People and Places:

Social Landscape Characterisation for inclusive and sustainable heritage management

Martina Tenzer

PhD

University of York

Archaeology

February 2024

Abstract

The historic environment—comprising a palimpsest of landscapes, buildings, and objects—carries meaning and is crucial in giving people a sense of place, identity and belonging. It represents a repository of ever-accumulating collective and individually held values—shared perceptions, experiences, life histories, beliefs, and traditions. These elements afford meaning-making, developing social values, and, subsequently, place attachment. Despite the Burra Charter and Faro Convention's aspiration to include people in the assessment process, individual, subjective, or emotional connections to place are often overlooked within heritage decision-making. Most changes to landscapes happen as part of the planning process, which is not currently able to account for individual connections but is based on views expressed in the language of the Authorised Heritage Discourse (AHD).

This research addressed the challenge to collect, analyse and meaningfully integrate individually held values into the assessment framework of heritage and landscape management. Artificial Intelligence methods, particularly Natural Language Processing and Topic Modelling, were successfully applied to survey, interview, and social media data to analyse the places and reasons behind the development of social values and place attachment. Categorisation, based on elements of *Grounded Theory*, and their visualisation have shown that individually held values form patterns of social values across wider landscapes. The people and place-centred method of *Social Landscape Characterisation* (SLC), resulting from this research, collects, analyses, and visualises these invisible or hidden value communities based on the same meaning (category value) or location (place value) as shared values across landscapes.

SLC provides a method for inclusive and transparent heritage and landscape management including people's individually held values in existing assessment frameworks of planning and decision-making. People-centred, place-based heritage and landscape management can increase the quality and distinctiveness of landscapes for managing the historic environment in a socially sustainable way for present and future generations.

Table of Contents

List of Figures

List of Tables

Author’s Declaration

I declare that this thesis is a presentation of original work for which I have primary responsibility. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged in the **References** section.

This thesis includes chapters which have been published in peer-reviewed journals (**Chapters 2-6**). For the co-authored papers presented as **Chapters 3, 4,** and **6,** I provide details of the respective contributions of each of the co-authors in the *Additional Information* section at the end of each paper. The included papers are:

Tenzer, M. (2022). Tweets in the Peak: Twitter Analysis - the impact of COVID-19 on cultural landscapes, *Internet Archaeology 59*. *DOI: 10.11141/IA.59.6*

Tenzer, M., & Schofield, J. (2024). Using Topic Modelling to Reassess Heritage Values from a People-centred Perspective: Applications from the North of England. *Cambridge Archaeological Journal* 34(1), 147-168. *DOI: 10.1017/S0959774323000203*

Tenzer, M., & Schofield, J. (2023). People and places: using Topic Modelling to identify reasons for place attachment – case studies from the north of England. *Landscape Research* 49(3), 340-358. *DOI: 0.1080/01426397.2023.2289970*

Tenzer, M. (2023). Social landscape characterisation: a people-centred, place-based approach to inclusive and transparent heritage and landscape management. *International Journal of Heritage Studies 30 (3)*, pp. 269-284. *DOI: 10.1080/13527258.2023.2289424*

Tenzer, M. et al. (2024). Debating AI in archaeology: applications, implications, and ethical considerations. Internet Archaeology 67. https://doi.org/10.11141/ia.67.8

Furthermore, I acknowledge that I received assistance from a professional proofreader, family member, and Grammarly to proofread **Chapters 1,2**, and **8** of this thesis in line with the Policy on Transparency in Authorship in PGR Programmes.

The included published papers in **Chapters 2** to **7** underwent copy editing at the respective journals in advance of publication.

Acknowledgements

Thanks to all participants in my surveys and interviews for taking the time to contribute to this research and share their personal experiences and connections to the places and landscapes that matter most to them.

Many thanks to Anna Badcock, Catherine Parker Heath (PDNPA) and Zac Nellist (SYAS) for their partnership and support throughout this project. With their help, I was able to find the people and their stories of places and landscapes.

Big thanks to my supervisors, Professor John Schofield and Professor Julian D. Richards, for their advice, support, constant encouragement, contributions, feedback and inspiration. This support has helped me develop my research skills and gain the confidence and courage to reach beyond my self-set limits.

Huge thanks – last but not least – to my wonderful husband, Torsten, for listening, encouragement, feedback, proofreading, unwavering support in times of uncertainty and – most of all – for believing in me.

Abbreviations

AHC	Australian Heritage Council (since 2004)
AHD	Authorised Heritage Discourse
AHRC	Arts and Humanities Research Council
AI	Artificial Intelligence
DTM	Document-Term-Matrix
ELC	European Landscape Convention
HE	Historic England (English Heritage until 2015)
HER	Historic Environment Records
HLA	Historic Landscape Assessment
HLC	Historic Landscape Characterisation
LCA	Landscape Character Assessment
LDA	Latent Dirichlet Allocation
ML	Machine Learning
NER	Named Entity Recognition
NLP	Natural Language Processing
PDNPA	Peak District National Park Authority
SA	Sentiment Analysis
SYAS	South Yorkshire Archaeological Services
TM	Topic Modelling
UTF-8	Unicode Transformation Format – 8 bit
OS	Ordnance Survey

Chapter 1:

Introduction

1.1 The study in context

Since ancient times, Greek and Roman cartographers have created maps for different purposes: to structure the world, give orientation, colonise or control new territories, navigate by sea or on land, or define areas of ownership and land use. In the mid-15th century, an extraordinary world map was drawn up by the Venetian monk and cartographer Fra Mauro (Brotton 2015, 75). This revolutionary map was created from earlier maps and historical documents and included spatially referenced written and oral histories of travellers, with over 3,000 descriptive texts artistically integrated (Figure 1-1).

Recently, this form of visualising the essence or character of a place has been revived as 'deep mapping' (Bodenhamer et al. 2015; Harris 2015) almost 500 years after Fra Mauro's masterpiece of cartography. Following form what is now referred to as a 'spatial turn' and a 'cultural turn' (Cosgrove 2004, 57; Earley-Spadoni 2017, 95; Jahn and Buchholz 2010, 511; Pendlebury and Gibson 2016, 1-2), this provides a tool to represent people's experiences, life stories, beliefs, traditions, and favourite places.

Figure 1-1: World map drawn by Fra Mauro (around 1450) including texts of travellers (source https://en.wikipedia.org/wiki/Fra_Mauro_map#/media/File:FraMauroDetailedMap.jpg)

Connections between people and places create an archive of local knowledge on everyday heritage that gives people meaning, a sense of identity and belonging. This sense of place can be translated as social values that constitute the quality of places.

Understanding such qualities of place provides a crucial background for planning and decision-making. This process has been gradually transferred from central to local governments, integrating local people's perceptions, visions, and needs and thereby creating places with qualities that benefit local community coherence and well-being (United Nations 1992). Additionally, national guidance and international charters and conventions encouraged the dialogue between experts and laypeople, drawing from local knowledge and empowering communities (e.g., Council of Europe 2000; 2005). At the same time, the focus of local policies shifted towards creating resilient communities and heritage, and the historic environment is increasingly seen to support well-being and community cohesion. Fostering a sense of place and belonging has advanced as a key principle in place research and in local planning and community initiatives (e.g., Cresswell 2015; Feld and Basso 1996; Jones and Leech 2015; Seamon 2020). Appreciation of heritage and connection to place was seen as positively creating place attachment (Lewicka 2011; Altman and Low 1992) and a sense of belonging and identity (Feld and Basso 1996; Graham, Mason and Newman 2009; Jones 2017; Madgin and Robson 2023; Nardi 2014; Tuan 1980). A deep connection to place as a form of 'Topophilia' (Tuan 1990) can include everyday and mundane places as well as designated heritage as defined by heritage experts. The aims and aspirations of government advisers, such as Historic England in the UK, and international bodies, such as the UNESCO ICOMOS and the Council of Europe, regarding landscapes and communities, have been influential and forward-looking. In the past 50 years the need to develop tools and methods for practical applications that put these ideas and aspirations into practice have led to a wide range of academic research projects (Social Value Toolkits), local government initiatives, community projects, and programmes from organisations such as Historic England (HLC, Conservation Principles and Everyday Heritage grants).

Historic Landscape Characterisation (HLC), developed by English Heritage[1] since the early 1990s (Historic England 2021) (see Figure 1-2 for an exemplary view of HLC), was developed to capture the character of a landscape based on the historic development of areas. This principle implies that all landscape is the product of human interaction and, therefore, culturally or socially constructed (Bradley et al. 2004, 6; Byrne, Brayshaw and Ireland 2003, 50; Byrne 2008, 155; Darvill 1999, 107; Phillips 2005, 20; Schofield 2014, 2). HLC was designed to support local authorities in planning and development as one component of a modular framework within the planning and development process (Clark, Darlington and Fairclough 2004). A further tool in the framework for planning and development control was created with Landscape Character Assessment (LCA), which provided more opportunities for the integration of local people's needs and values (Tudor 2014). These expert-led methods aspired to include experiences and attachment of the communities, integrating local knowledge as a background for local planning in a proactive way. Several projects detailed and discussed later in this chapter successfully engaged local communities and contributed to positive change and development of places. However, the meaningful integration of social value – of 'soft', fuzzy, inherently subjective data – in (apparently) objective, fact-based heritage data sets that fed the decision-making process in local planning and define significance in designation processes for heritage assets has long been seen as challenging, if not impossible (Dalglish and Leslie 2016, 217).

This thesis will propose an innovative method for social value analysis and situate this approach within the wider research context. This research will present social value mapping of individually held values on a landscape scale, which can feed into existing social value projects that are commonly site and group focussed.

Firstly, this thesis will provide an overview of current approaches to and tools for social value and place attachment assessment. Through discussion and critical review of existing tools and methods for social value, research gaps will be identified, for which this thesis will provide solutions. Changing perspectives and attitudes in

[1] English Heritage split into Historic England and English Heritage in 2015. Research and consultancy, including the HLC project, are now part of Historic England's responsibilities.

heritage discourse and social value research over the past decades will provide a wider context for this research. Furthermore, the topic of categorisation, narrative approaches and mapping of social values will be elaborated on.

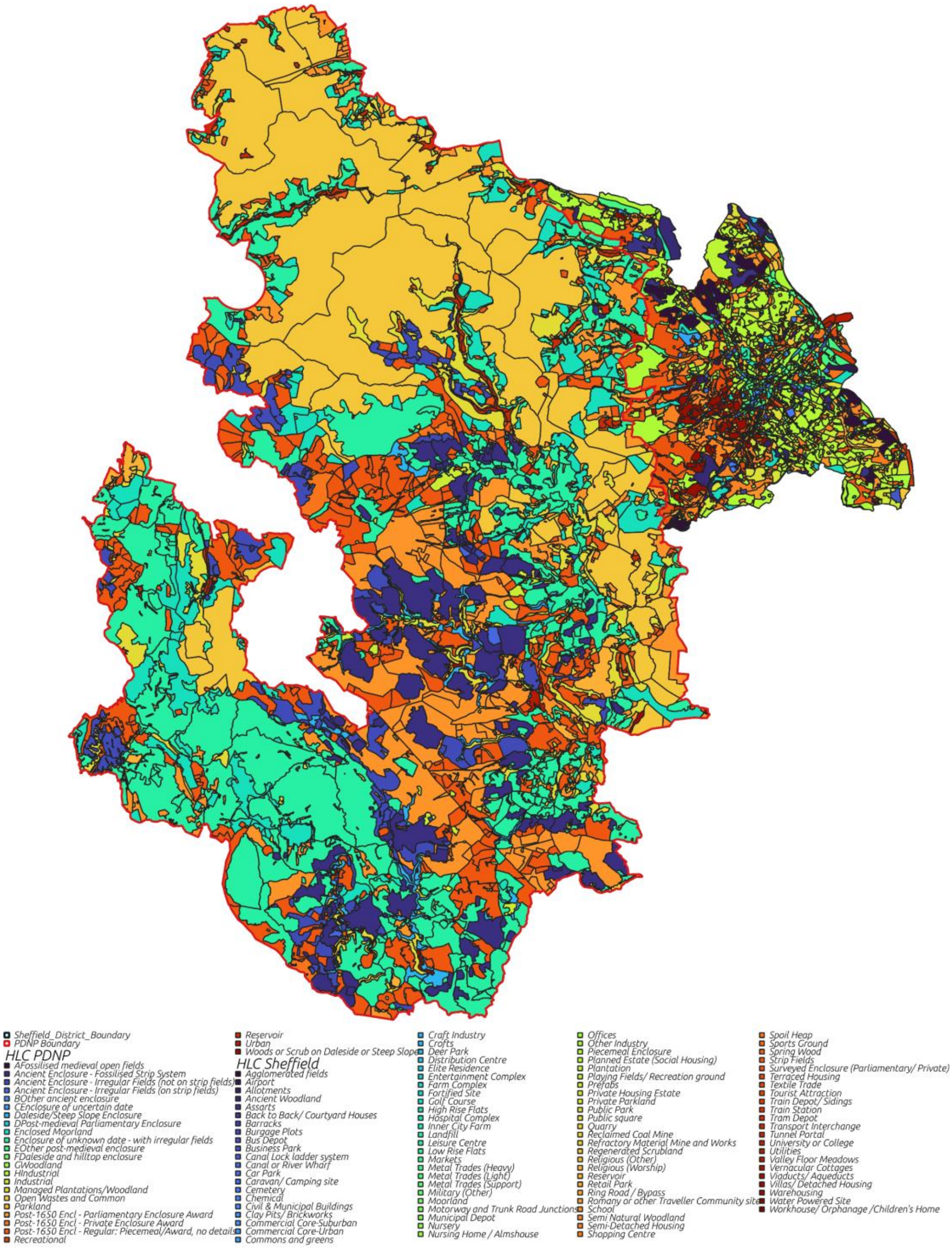

Figure 1-2: HLC map of the Peak District National Park Authority and Sheffield City (map created in QGIS, data by ADS).

Secondly, this thesis will offer a methodology towards understanding and visual representation of people's 'stories' and personal connection with places by defining *social values* as reasons behind a strong place attachment to living and working environments. The methodology is based on adopting and adapting AI techniques and translating the results to spatial data visualised in digital maps. One aspect of this thesis will focus on the definition and influence of 'everyday' heritage shaping working and living environments and landscapes and explore what role it plays in creating place attachment.

Thirdly, this research will demonstrate how this data can be translated a layer in socially sustainable heritage and landscape assessment frameworks, representing crucial background knowledge for local authorities to avoid tensions and conflict with local communities in the decision-making process. Finally, practical mapping will demonstrate how this methodology can feed into a guideline for socially sustainable, proactive planning and demonstrate its potential in practice to promote the use of this tool for inclusive decision-making. This method has the potential to enhance dialogue between local governments and the public and identify highly valued places on a landscape scale and in different environments, aiding a better understanding of people-place connections. This crucial background enables proactive planning for socially sustainable heritage and landscape management, enhancing the quality of places and strengthening people's sense of place, belonging, and identity.

1.2 Review – social value assessment, mapping, and toolkits

1.2.1 Introduction

The first projects focussing on the connection between people and places through their perception of the environment developed in the second half of the 20th century. Over the past two to three decades, this research has intensified for public benefit and to address the aspirations of international agendas and treaties on sustainability (see **Chapters 2.1** and **2.2**). Local agendas and initiatives gained traction through integrating participation and consultation with local communities in their planning and decision-making. Academic research institutions, national and local governments, and various organisations and charities produced considerable research outputs, from funded research projects to participatory activities and toolkits or guidelines.

The following will provide a comprehensive overview of past and current projects aiming to assess and/or map people-place connections, place attachment, landscape qualities, landscape perception, and preferences. The approaches taken in earlier projects will be discussed, and a gap will be identified for which this thesis will provide a solution.

1.2.2 Community and counter mapping

Maps are a familiar medium even to people without deeper knowledge of cartography (Perkins 2007, 127), and creative mapping offers a way to communicate the experiences of laypeople. Maps are an ideal means to share views and perceptions (Fairclough 2002, 284; Wood 1992, 79); they are a familiar way of representing and orientating oneself in the world that most people understand intuitively (Clifford and King 1996, 5; Perkins 2007, 127). Maps can be more than a cartographic expression or modelling of the world. The usual form of maps represents the world as an abstract model, generalised and focussed on supporting specific and sometimes very narrow questions. These are referred to as 'thin maps' (Bodenhamer et al. 2015; Harris 2015).

Community mapping and counter mapping projects engaging the public with their places developed decades before the Landscape Convention adopted this concept. Early projects, combining public engagement, participatory methods and people's perception, focussed first on urban environments. For instance, Lynch's cognitive maps of city dwellers in Boston and Los Angeles in the 1960s were a remarkable example of how people conceptualise and structure the environment of everyday life, creating maps of daily routines (Lynch 1960; 1972). It emphasised the discrepancies between the professional view of experts in planning and the perception and interpretation of laypeople in everyday situations. Since then, academics and practitioners have explored the different pathways that allow public participation and inclusion of local views, and experiences and place attachments in guidance and legislation.

Community mapping can take many forms, from artworks to 3D models, as demonstrated in the following examples. Common Ground was an environmental and arts organisation in the UK that led various community initiatives to express local perspectives on place and character[33]. Outcomes of the projects, for example, the *Parish Maps* projects and the *Alphabet of Distinctiveness,* were artistic articulations of a sense of place and belonging (Common Ground 1996; 2006). However, neither project was adopted in the characterisation process itself or used for practical applications in the planning process.

Di Nardi's work (2014) co-created an experiental 2D paper map with the local community to present a counterpoint to official maps used by heritage professionals and provide insights into how the local population valued their local area. Similarly, a community project in the village of Slaithwaithe, UK, used a 3D model of the area to assess people's connection to place (Craig, Harris and Weiner 2002). Low-tech methods have the advantage of enabling the participation of a wider range of people without the demand for pre-existing computer skills or technical knowledge and reducing the time for 'skilling-up' and workshops (see **Chapter 1.5.4** and Dabaut 2021, 254). However, non-digital project outputs are challenging to integrate into a framework of existing digital databases and maps.

The notion of spatial perception and cartography of experience and practice led to projects such as *Mapping Attachment* (Byrne and Nugent 2004; Byrne 2008a, 2014; Harrison 2011; Perkins 2007) or *Bio mapping* of the emotional attachment of people to places (Nold 2009) which developed capabilities through GIS by introducing a layer of meaning in a spatial system (Perkins 2007, 128). Counter mapping developed into an essential tool for social value assessment (Byrne 2008a, 2014; Harrison 2011; Schofield 2014). This form of representation of meaning, feeling, experience and perception – intangible spatial aspects – is based on subjective, 'soft' and often 'fuzzy', and possibly non-spatially referenced data (Craig, Harris and Weiner 2002, 111). Another example of this approach was realised in *Proboscis*, a multidisciplinary, multi-organisational project set up by two London-based artists[34]. In an approach of 'co-discovery of

---

[33] See also **Chapter 5:** Social Landscape Characterisation: A People-Centred, Place-Based Approach to Inclusive and Transparent Heritage and Landscape Management.

[34] See http://proboscis.org.uk.

uncommon insight', the project aimed to collect and represent data through 'social engagement, creative research, innovation and problem solving'. The deployment of GPS and other geographical referencing devices, such as mobile phones, and their integration into a GIS, enabled cognitive mapping of lived experiences similar to Lynch's early cognitive mapping approaches.

Kiddey's project (2014) assessing the place attachment of homeless people in Bristol and York proved that connection to and rootedness in place can also develop among groups with no permanent home (see **Chapter 2.5.** for different views on place attachment development).

Currie and Correa (2022) provided a method for mapping tangible and intangible elements of Edinburgh's culture, focussing on the event and culture sector.

The examples presented in this section show the wide variety of approaches mainly focussed on communities and urban environments. Because 2D maps usually represent snapshots in time, their usefulness for the representation of heritage or cultural aspects of landscapes has been questioned due to the dynamic and perception-based nature of these categories (Rudolff 2006; Smith 2006, 71, 80)[35]. (Golledge 2006). The quantitative spatial sciences' ability to deal with the challenges of the time and have 'social relevance', politics and power was questioned (Cox 2014, 54-55). National parks have been targeted to assess people's sense of landscape quality, as well as the next section will show.

1.2.3 Public participation in national parks

National parks have been targeted for qualitative research and ethnographic studies over the past decades. The closed bounded area with larger parcels of similar character and specific qualities seem to afford aspects particularly suitable for this kind of study. Key research questions in national parks range from park improvements to environmental aspects. For example, Brown and Weber (2011) describe a method for Public Participation GIS (PPGIS) to combine visitor perception, required facilities and

[35] However, maps can be biased, misinterpreted and misused (Herring 2009, 70; Monmonier 1996); similarly they disempower people who are omitted from them (Byrne 2008, 256).

environmental impact for national park planning in the Greater Alpine region of Victoria, Australia. National Park research also aims to understand the pro-environmental attitudes of visitors and residents[36] (Petrova, Čihař and Bouzarovski 2011; Hausmann et al. 2020; Ramkissoon, Weiler and Smith 2012).

Of particular interest to this research concerning people's perception and attachment to places are ethnographic studies conducted in national parks. For example, Taplin, Scheld and Low (2002) focussed on the Independence National Historical Park in their Rapid Ethnographic Assessment Programme (REAP), which provides a set of qualitative methods to assess people's connection to places.

Dabaut (2021) undertook a study in the Northumberland National Park in relation to HLC (see below). Similarly, Maguire (2017) focussed on a national park when assessing the perceived qualities and place attachment in the Colliery Dam Park, British Columbia, Canada to create a landscape preference map.

Further large-scale ethnographic research to inform the national park management, was, for example, undertaken by the National Park Service in the US[37]. These programmes use the developed strategies of REAP, oral and life histories and ethnographic landscape studies to improve visitor experience, strengthen historic relationships and traditional use of parks, and to inform a better park management. While the studies collect social value in the study areas, they do not attempt to categorise the character of and attachment to the landscapes based on social values.

1.2.4 Mapping landscape attractiveness and place attachment

Particularly in the field of eco services, mapping of landscape quality and preferences has led to a variety of mapping approaches, for example, in relation to urban woodland in Helsinki providing insights into landscape quality based on scoring landscape types (Tyrväinen, Mäkinen and Schipperijn 2007). Also, PGIS was used in a

---

[36] This is of particular interest since national parks in the UK play an important role in contributing to the 30 by 30 target to boost biodiversity and nature recovery. (https://naturalengland.blog.gov.uk/2023/12/11/30-by-30-a-boost-for-nature-recovery/)

[37] https://www.nps.gov/ethnography/parks/approaches/index.htm

research that focussed on forest planning and people's landscape perception in resource management to understand how landscape values influence the reasons behind the development of preferences (Brown 2004).

In a more extended landscape approach, De Vries (2007; de Vries et al. 2013) focussed on mapping the attractiveness of landscapes in the Netherlands and asked participants to assess the qualities of six study areas based on a predefined scoring system. The study will provide a method that can feed people's perceptions into the process of impact assessments and cost-benefit analysis. In another example, Cinderby et al. (Cinderby, Snell and Forrester 2008; Cinderby et al. 2012) mapped environmental qualities to improve conditions of the lives of residents in urban spaces.

Place attachment research was particularly focussed on developing methods for measuring and mapping connections between people and places over the past decades. Attachment GIS maps and scoring systems were created based on surveys, mainly using Likert scales and interviews (Brown and Weber 2011; Boley et al. 2021; Brown, Raymond and Corcoran 2015; Brown and Raymond 2007; Maguire 2017; Scannell and Gifford 2010).

Studies focussing on aspects of the natural environment in correlation with people's perception contributed to a better understanding of ecosystems in relation to people's connection to places. This category of studies provides a good overview of tools for mixed-method approaches (e.g., Likert scales and qualitative questioning) and visual representation techniques. However, as these approaches commonly use predefined landscape value categories or structured questioning, they lack the deeper insights of narrative approaches. The latter provides a deeper understanding of people's individual connection to places that form social values.

As a disadvantage of participatory community projects, or *Participatory GIS,* they risk of being bound into an agenda of experts or researchers who function as facilitators in community mapping projects or, in another extreme, uncover deep-lying tensions in a community and draw aggression against the facilitator (McGhee 2012). In particular, *Participatory GIS*, which relies on complex technology and software, has been seen as problematic for the use of non-experts (Poplin 2012) and because of bias towards specific research agendas (Perkins 2007, 127).

1.2.5 *Deep mapping and lived experiences*

The step from 2D maps (see Nardi 2014) or physical 3D models (see Craig, Harris and Weiner 2002, 133), presented in **Chapter 1.2.2**, to deep maps produced in GIS[39] offer new opportunities to social value mapping. Advances in digital technology and GIS triggered the recent 'spatial turn' in digital humanities (Cosgrove 2004, 57; Earley-Spadoni 2017, 95; Jahn and Buchholz 2010, 511) and supported projects exploring emotional, experiential, and phenomenological aspects of everyday lives and lived-in worlds. It found its latest expression in 'deep maps' which were developed to finally convey the 'thick descriptions' originally envisaged by Geertz (1973) allowing the integration of subjective, descriptive, and text-based data into GIS (Harris 2015; Bodenhamer et al. 2015; Earley-Spadoni 2017, 96-97; Kwan and Ding 2008).

Deep maps could be seen as the technological descendants of Fra Mauro's map with stories and cartographic detail in one plane, towards developing layers of stories, images and other media. Also, in contrast to the one-plane map, deep maps have the potential to display various, even contradicting information for the same location and, therefore, 'allow[s] for dissent and discussion of contested geographies, and furthermore permit multi-vocality' (Earley-Spadoni 2017, 97).

Deep maps can lay out maps of abundant, diverse information derived from various media to create a deeper understanding of the social fabric of a landscape – of stories, history, and lived experiences. This is demonstrated, for instance, in the RICHES project of the University of Central Florida[40], which provides a platform to compile local historical data with other datasets and create their individual narrative of places (Earley-Spadoni 2017, 96-97). As a further example, the LANDMAP project carried out by *Natural Resources Wales* created a methodology to map cultural patterns across landscapes in Wales[41]. The map contains aspects of what constitutes

[39] Great advances in computer capacities and capabilities allowed the development of the Geographic Information Systems (GIS), which allows fast, cost-effective adaptation, automatic import, and analysis of new data.

[40] https://riches.cah.ucf.edu

[41] http://lle.gov.wales/catalogue/item/LandmapCulturalLandscape?lang=en

the local identities of communities. This project was again based on the work and assessment of heritage professionals (Natural Resources Wales 2016a; 2016b). Particularly of interest for this research is the attempt to create a '"typology" of cultural landscapes' (Natural Resources Wales 2016, 7).

A project with a particular focus on ageing and connectivity, representing the deep mapping of life stories and experiences, was undertaken by the University of West England. The project 'Either side of Delphi Bridge' (Bailey and Biggs 2012) visualised the connection between local elderly residents and the social and cultural fabric of their place of residence in digital form.

The UNESCO-supported *Cultural mapping* project provides a new method of mapping intangible and tangible cultural aspects of places[42] (Cook and Taylor 2013). The guidelines recommend a series of techniques for knowledge mapping, such as topic maps, cognitive maps, and mind maps (Cook and Taylor 2013, 185-242). Several academic and commercial projects based on the project's methodology achieved the map-based representation of abstract concepts, such as cultural identity across various urban and rural areas (see Currie and Correa 2021; Currie and Correa 2022; McKeithen 2015).

Further examples of place-based value research will be discussed in **Chapter 5**. This will also include the AHRC-funded place research project *Place matters: the arts and humanities and the place agenda*[43], which included a number of academic projects and local initiatives (Madgin and Robson 2023). For example, 'Roots and Futures' [44], a map-based representation of areas in Sheffield, combined historical and archaeological data with the option for local residents to add their memories and experiences to the project.

Similar approaches are used in a series of local government initiatives as part of the 'Know your place' projects [45]. The GIS dataset provides, among other information,

---

[42] https://unesdoc.unesco.org/ark:/48223/pf0000159090

[43] https://www.ukri.org/blog/place-matters-the-arts-and-humanities-and-the-place-agenda/

[44] https://www.sheffield.ac.uk/archaeology/research/roots-and-futures

[45] See, for example, the West of England map: https://www.kypwest.org.uk/

layers of historic maps, photographs, and an interface for community contributions. It also allows the community to add walks and special places to the local authority platform.

Approaches that combine historical place data with individual connections of local residents provide a deeper insight into people's connections to places. In the past, projects focussed on creating deep maps with layers of different information and integrating the information of local residents to develop a comprehensive picture of places. Projects such as these have focussed predominantly on urban areas and dependent on people's active registration and participation in online websites for mapping. Barriers to online participation should be considered when using this format of co-creation and as means of communication between the public and heritage professionals.

1.2.6 HLC and social mapping in policy and management

The aspiration of the HLC development team from the outset was to inspire and integrate people's perceptions and experiences through the historic landscape character maps[46]. This idea was promoted by Turner (2007, 46) as one use of HLC: to 'help people recognise and create new narratives and ideas about their landscapes'.

The latest HLC was a project in Oxfordshire that attempted to engage the public with a series of talks and events, inviting communities to express their view on the landscape's characteristics (Tompkins 2017, 463-465). However, the outreach activities resembled more a tick-list exercise for including public opinion than an effort to include the information gathered into the catalogue of resources from which the final character map was developed. As in other projects, the sources for creating the character were based on the usual data which provides evidential and historical information. The resources listed in the project report do not mention public opinion or participation as a source of information (Tompkins 2017, 19). However, the HLC project was supposed to 'provide data, by which individuals, community groups, or

---

[46] For a detailed description of the method see **Chapter 2.6.2**.

academics could research and engage with the historic landscape.' The additional 'Living landscape' project of the Oxfordshire HLC created a treasure trove of local knowledge consisting of drawings, paintings, historic walks and 'post-it poetry'; however, the opportunity to include this data meaningfully into the HLC dataset was missed as this project was only intending to 'increase awareness' rather than include people's views and perspectives (Tompkins 2017, 447).

How the meaningful integration of local people's landscape perception and values can be undertaken is shown in the guidelines of the charity Campaign to Protect Rural England for the creation of community landscape character assessments (Campaign to Protect Rural England 2018a; 2018b). The guide provides an HLA method based on local community expertise that can subsequently be meaningfully integrated into local authority's planning systems.

Several UK-based and international projects applied the HLC method for public benefit or inclusion of social values in the assessment. For example, Dabaut (2021) based his participatory approach to assess people's connection to landscapes in the Northumberland National Park. He used walking interviews and surveys as data collection methods. The approach provided a detailed insight into a small area of the national park. While the preferred method was identified as the walking interview technique, such an approach cannot be realised on a landscape-scale because of the time-consuming nature of this tool.

Another example, how HLC can help manage urban green spaces for public benefit was proposed by Dobson and Selman (2012). In cooperation with local authorities, they connected green spaces in Sheffield to enhance urban planning and management. Internationally, examples of HLC use are shown in the terraced agricultural landscape in the Mediterranean (Turner 2018, 47), such as in Turkey (Turner and Crow 2009). The application of the method revealed a deeper historic time-depth and more complex development of the landscapes in the past than anticipated. Also, Gaffney and Dingwall (2007) used the general approach of HLC for a project in Fort Hood, Texas to enhance the method by acknowledging that past events, such as the enforced movement of the population from the military base, affected people's perception (Gaffney and Dingwall 2007, 1). However, while the project used archival material on oral histories and conducted outreach, no integration of people's current perceptions appears in the final assessment.

More extensive projects on a grassroots level have been described by Dalglish and Leslie (2016, 217-224) working with a community in Govan, a part of the City of Glasgow with a long history and shipbuilding tradition. In a surge of activities and direct engagement with planning decisions, the community was able to react to change and development in their neighbourhood. The initiatives often developed in 'a chaotic way' but through 'hard work' by the local diverse and heterogeneous community and community council (Dalglish and Leslie 2016, 219), funding was redirected and the quality of their place was enhanced through an 'unmanaged and ongoing process of characterisation' (Dalglish and Leslie 2016, 219). Although in contrast to the Common Ground Parish Map projects, this example directly impacted policies, funding, and decision-making, such initiatives are not replicable and challenging to conduct for local authorities because of the technology and resource input necessary to facilitate such projects. This approach was reactive to planning decisions or a 'problem-orientated characterisation process' (Dalglish and Leslie 2016, 224). This enormous effort by the community is not a template for others and 'could hardly be replicated' as it emerged from the community's desperate situation (Dalglish and Leslie 2016, 224), but it might be an inspiration for more practical methodologies. A method to capture community views and attachment on an ongoing basis would be more effective and efficient, as well as practical within budget, personnel, and time constraints. It could be accessible online and proactive, in advance and as background for the decision-making process. Primdahl and Kristensen provided an example of community integration in the planning process. However, the study proved to be highly technology-driven and, with qualitative methods such as focus groups, the practicality of this approach in the daily work of local authorities would be a challenge. There is also the issue of privileging groups and community members that are vocal and dominant and missing voices of underrepresented parts of the community (Jones and Leech 2015, 30).

While HLC is routinely used with other data sets and maps (Herring 2009, 75), it is challenging to include social values into the framework of characterisation. However, these forms of representation and databases of subjective, 'soft', and phenomenological data associated with other spatial heritage or landscape management information, such as Historic Environment Records (HER)[47] and National Character Areas (LCA)[48] data, would bridge the gap between the traditional expert knowledge and the local knowledge of everyday lived-in worlds. The approaches show that HLC has the potential to provide a basis for the integration of people's perceptions and social value assessment. However, the study areas of these projects were site-based and focussed on a small study area or a particular element of a larger area. None of the approaches aimed to understand and map social values on a landscape scale or produce a methodology that would have the potential to be scaled up to cover wider landscapes, which is one of the key principles of HLC. A deep map platform for landscape characterisation not only representing 'official' facts and data, but also integrating public perceptions that diverge from the official picture presented by HLC, would offer an opportunity to give a holistic representation. These will be developed in this research, creating a form of Social Landscape Characterisation (SLC), which characterises and visualises the social aspects of landscapes individually. Such SLC maps could be an essential background for local authorities' planning decisions and enable sustainable change and development[49].

1.2.7 Social value assessment toolkits

Toolkits and guides equip communities and groups with a tool to express and systematically record social values for meaningful integration in planning and landscape management. Over the past decades several such toolkits were developed to enable the communication between local people and authorities. For example, the HLA toolkit of the *Campaign to Protect Rural England*, described above, is one such toolkits that provides a framework for character assessment. Dalglish and

---

[47] https://www.heritagegateway.org.uk/gateway/chr/

[48] https://nationalcharacterareas.co.uk/

[49] See **Chapter 5**.

Inherit/Community Land Scotland (2018) identified a gap in public participation in relation to landscape management and developed a framework for the integration of public perception and local knowledge in the planning and management process in Scotland.

Another compilation of proven methods for social value assessment on a community basis was developed by Robson into a *Social Value Toolkit*[50] launched in 2021. It presented a wide range of methods to assess social values in relation to heritage. The toolkit was aimed at heritage practitioners to enable local social value assessment in cooperation with local people, communities, and groups. The project provided six case studies that applied different data collection methods, such as interviews, observation, and counter mapping.

The Chartered Institute for Archaeologist's public Engagement Toolkit[51] provides instructions and advice to engage the public with archaeology actively. The project aims to understand better how archaeology can contribute to public benefit and social value. Other toolkits that aim to work towards the socially sustainable transformation of urban places are, for example, the European Deep Cities programme 'CURBATHERI'[52], a collaborative project of the Universities of Florence, UCL, University of Stirling, the Norwegian Institute for Cultural Heritage (NIKU) and the University of Barcelona (for the Deep Cities project of the University of Stirling see Jones et al. 2024)[53]. Another example of a social value toolkit focussing on architecture and the benefits of understanding social value in the urban context is provided by the

[50] https://socialvalue.stir.ac.uk/

[51] https://www.archaeologists.net/toolkits/community-archaeology/1-1_archaeology-public-engagement

[52] https://curbatheri.niku.no/

[53] https://www.deepcities-toolbox.unifi.it/p21.html

University of Reading [54]. This is part of the wider project of 'Social Value UK'[55], a professional body for social value and impact management.

These tools and guidelines focus on a community and site level. Such methods provide the basis for reactive community interventions, similar to projects of community characterisation presented in **Chapter 1.2.6**. A method that allowed to identify interest or value communities and places with high value in advance of change and development would provide a background which similar projects, as described in this chapter could build on.

1.2.8 Conclusion

The examples explored in this chapter show variable degrees of participation of communities and groups, as described in Arnstein's Ladder of participation (Arnstein 1969, Fig. 2). These range from creative expressions to Participatory GIS – from passive consumers and educated masses, such as visitors to museums and informed communities, to a consultation process in which communities could voice their visions, needs and aspirations. Participation can reach as far as a meaningful partnership and dialogue between communities and local authorities, which proved highly influential in the planning process, as shown in the examples of the Danish case studies or the Scottish initiative (see **Chapter 1.2.6**). Furthermore, guidelines offer communities ways to influence local planning policies through input of local knowledge into, for example, neighbourhood plans or village design statements (Campaign to Protect Rural England 2018; Clark, Darlington and Fairclough 2004, 52). These guides offer potential opportunities to increase and include the idea of a 'sense of place' in the wider framework of heritage assessment and landscape characterisation from a people's viewpoint. The number of toolkits and guidelines for social value assessment and community mapping, focussing on various aspects of the historic environment, have

[54] https://www.architecture.com/knowledge-and-resources/resources-landing-page/social-value-toolkit-for-architecture

[55] https://socialvalueuk.org/value-toolkit/

increased over time and several of these are available and accessible online at the time of writing (see **Chapter 1.2.7**).

The next section will discuss these approaches regarding their strength and weaknesses and identify the gap which this research will address.

1.3 Discussion on current social values approaches and research gaps

This chapter has introduced a number of projects, initiatives, and guidelines for social value and place attachment assessment providing practical methods and tools for a better-informed management of places. The examples presented in this research use similar methods for data collection, for instance, interviews, surveys, mapping, observation, focus groups, and social media. As Johnston (2023) pointed out, the current focus of social value assessment is on groups or communities. Community work, such as focus groups, can be biased towards dominant voices in a group and negotiation of common values that may not represent individual opinions . Communities participating in research or local planning forums often consist of self-selected, active community members who dominate the decision-making process, which can obscure underlying opinions of less vocal community members (Craig, Harris and Weiner 2002, 101; Dalglish 2018, 55-58; Jones and Leech 2015, 30). Where individuals were involved, for instance, in the individual interviews and observations of Robson's Social Value Toolkit (see **Chapter 1.2.7**), these were mainly members of predefined communities or groups and related to predefined heritage assets or places[59]. Similarly, the *Everyday Heritage* projects of Historic England aim at predefined sites and groups identified by heritage experts. Larger mapping exercises focussed on place attachment (Brown and Raymond 2007) and environmental preferences (Maguire 2017). Community projects tend to be reactive in relation to change and development in the planning process (see **Chapters 1.2.6** and **1.2.7**, and

---

[59] See, for example, the case study of Cables Wynd house, one of the six case studies of the toolkit, available at https://socialvalue.stir.ac.uk/files/2021/01/Site-Report-Cables-Wynd-House.pdf

selected case studies in the *Social Value Toolkit*[60]).

As shown, present approaches rely on either predefined groups, the assessment of landscape qualities, or social values in small-scale landscape approaches. There is, therefore, a need to understand social values a) on an individual basis beyond communities and groups that have a collective interest in a place and b) to represent individually held values on a landscape scale that allows identification of previously unknown commonalities in social value across wider landscapes – rural or urban. A combination of current approaches, while integrating social value assessment of individual people into landscape-scale mapping, would provide a dataset that could be updated to address the dynamic and fluid nature of such values. This would offer a tool for correlating with current heritage datasets. Such an approach would enable proactive planning and development and equip local planning authorities with essential background knowledge of social value.

Furthermore, Currie and Correa (2022, 101-102), in their study on Edinburgh's cultural landscape, concluded that 'codifying this subjective information simultaneously strips it of the richer narrative participants told about place'. This challenge to codify or categorise the narratives of local people remains problematic in the narrative approach to social value assessment (see **Chapter 2.4** for the categorisation dilemma). This research will provide a solution to qualitative categorisation using Artificial Intelligence in the process of qualitative data analysis and propose a methodology for narratives in thematic analysis.

The research gap identified can be summarised as a need for a methodology that combines traditional qualitative research methods, e.g., interviews, surveys, and social media data, with AI tools to identify individually held social values and represent these as patterns on a landscape scale. This method could then provide the starting point for identifying social value hotspots as the basis for applying existing social value assessment tools. For example, the *Social Value Toolkit,* with its focus on communities interested in a particular site that is subject to local planning, can identify specific individuals with common social values in a place or understand the values in a

[60] https://socialvalue.stir.ac.uk/case-studies/

particular place from this research. In an urban environment, such as the *Deep Cities* project, the data set resulting from this research could help identify areas or objects of higher community interest to undertake targeted research that gains a deeper understanding of the underlying connections and attachments.

The following sections will define the research questions and provide an overview of the methodology developed and applied in this research.

1.4 Research questions

The study's main aim is to explore opportunities for integrating local people's individual opinions, needs, and visions and develop a *Social Landscape Characterisation* based on the social values people hold in their everyday living and working environments, based on the heritage and historic landscapes in which life unfolds.

Specific objectives included:

- Objective 1: How can people's individually held values and the reasons behind these values ('soft' or subjective data) be collected, allowing a categorisation based on latent themes within the data, and analysed, using freely available and open-source software and code?
- Objective 2: Can Historic Landscape Characterisation or its key principles be adopted and adapted to accommodate people's perceptions and opinions on their living and working landscapes?
- Objective 3: How can social values be visually represented to create outputs for assessment frameworks within the planning and decision-making process and, at the same time, provide opportunities for developing engaging resources to increase participation for inclusive, transparent, and socially sustainable heritage and landscape management? The aim of this objective is to produce a guideline or methodology that can find practical applications in real-world scenarios.

Based on these objectives, the research focussed on collecting, analysing, and visualising the data. The methodology, as an overview, will be presented in the next section. Detailed methods and workflows are provided in the specific chapters pertaining to the respective publications, which comprise the thesis's body.

1.5 Study approach and summary of methodology

1.5.1 Introduction

This chapter will provide an overview of the methodology developed and applied in this research. The specific methods and tools are detailed in the published papers that form the body of this thesis and present the results of the analyses. The following overview will detail the rationale behind the overarching methodology, provide insights into how COVID-19 changed the initial approach, introduce tools applied for data collection and analysis, explain how study areas were identified and partnerships established, and describe sampling strategies and ethical implications.

1.5.2 The impact of COVID-19 and resulting adjustments

As a reaction to the larger-than-anticipated datasets that were the basis for this research, gathered from social media channels and surveys, I decided to change my approach to the data analysis from purely manual and NVivo analysis to Artificial Intelligence tools. On the one hand, this was partly because of the preference for open-source software in this research, which excluded the use of licensed software such as NVivo. On the other hand, it was because of the larger-than-anticipated dataset sizes, and the current advantages of AI in research and development for real-world applications. The opportunities and capabilities of this technology will be elaborated on in the publications in **Chapters 2 to 5**. Here, I will briefly introduce the technology and its benefits for analysing unstructured textual data.

The methods and approaches taken in this research, conducted between October 2020 and October 2023, should be seen against the background of the COVID-19 pandemic and related restrictions regarding social contact and free movement. While the circumstances did not negatively impact the research itself, it led to a rethinking and redesign of the methods, particularly in view of data collection adhering to social distancing during the lockdown phases of the pandemic. This reorientation in data collection, which involved retreating to remote and online practices, also provided different dataset sizes than initially anticipated. Social media data and online surveys provided larger datasets that required efficient and effective tools for analysis. The ethical implications, limitations and potential biases in the methodology will be discussed in **Chapter 6**. In the remainder of this chapter, I will provide an overview of the tools and methods developed and used in this research. The specific workflows applied to the different data sources will be elaborated on in the respective chapters as part of the publications forming the body of this thesis (**Chapters 2 to 5**).

1.5.3 Case study, area, and partner selection

The rationale behind selecting my study areas was mainly influenced by the previous partnership with the Peak District National Park for my MSc in Applied Landscape Archaeology and previous work as a commercial archaeologist in this area[62]. I also conducted archaeological work in the area of the City of Sheffield and had a prior connection with the staff of SYAS. This long connection gave me the advantage that I was already familiar with the procedures and staff at the partner organisations and the landscapes selected for case studies. Furthermore, the areas differ in their rural versus urban character, the population density and cultural offers, and in the character of the historic environment. The details of the study areas are described in the publications included in this thesis as **Chapters 2 to 5**. The definition of the study area also included the focus on target groups for participation in this research, which comprised visitors, local residents, and people working in the study areas.

As seen in **Chapter 1.2.3**, various projects have focussed on national parks to assess landscape qualities and social values. National parks have a range of advantages for landscape research; for example, the clear boundaries of the area can be used as a basis for the case studies, allowing for the understanding of personal connections between people and the parks. As such, a national park has one responsible park authority, which provides opportunities for cooperation and partnership. Another

[62] This PhD was not a Collaborative Doctoral Partnership, but the close partnership allowed me to benefit from existing structures and relations.

advantage is the availability of park-specific landscape and heritage datasets, e.g., the HLC dataset for the PDNP that provided a source for data correlation. National parks afford specific characteristics and qualities that are perceived as positive and beneficial for recreation. Additionally, the landscapes are less fine-grained. Urban areas represented in HLC are very fine-grained and can be challenging when assessed on a landscape scale. The comparison to the urban environment of Sheffield has shown challenges of HLC correlation on a landscape-scale and offers itself more readily to smaller-scale or site-based approaches. Nevertheless, the landscape-scale social value assessment works also in urban areas and can provide the basis for more in-depth approaches, such as those presented in the *Deep Cities* project and for the *Social Value Toolkit* (see **Chapter 1.2.7**).

1.5.4 Rationale for overarching methodology

The methodology of this research consisted of three steps: (1) data collection from three different data sources associated with the two study areas, (2) data analysis using Artificial Intelligence (AI) tools, and (3) data visualisation for the purpose of practical application of the results.

Categorisation of social values has been a challenge in previous approaches to social value assessment and mapping (see **Chapter 2.4**). This thesis will provide a solution for the difficulties surrounding qualitative data analysis, preserve the depth of narratives, and offer a method for the spatial interpretation and visual representation of social values across wider landscapes.

For two reasons, I decided against the commonly used qualitative analysis software NVivo (Welsh 2002)[63], as anticipated in the outset of this research. Firstly, the premise of this research lies firmly on the use of non-proprietary, freely accessible, and applicable software, data resources and code. Secondly, the opportunities afforded by AI tools, such as Topic Modelling, are an innovative, emerging technique that merits a greater acknowledgement and deployment within the heritage sector. NVivo has

[63] https://lumivero.com/products/nvivo/

integrated this technique into its ‘autocoding’ tool since Version 11 (update 2) in 2016, providing the ‘auto code wizard’ as an easy user interface and promoting the tool as an opportunity to get results that ‘are not influenced by your own bias’ as the automated insights are ‘computer-driven’ [64]. The issue with this approach was the ‘black box’ that this tool constitutes (for details on the ‘black box’ effect and explainable AI, see **Chapter 7**). To have the widest possible oversight of this process, the approach taken in this research used existing code and adapted this to serve the aim of this project. This approach enabled an open and investigative use of the Topic Modelling tool, as opposed to the proprietary TM software and parameters used by NVivo that are not specified in more detail. At the same time, this limited the wider adoption of this method by, e.g., local authorities or community groups because of the required basic coding knowledge to apply the tool. Nevertheless, the advantage of explainablity in the analysis and the security to perform all analyses without the need to store data externally was seen as an advantage.

AI is an evolving technology that is only slowly finding its way into the field of archaeology and Cultural Heritage Management (Traviglia 2022). While the technology has been successfully deployed for image and object recognition and classification (see **Chapter 7** for more detail on the applications), text-based analysis using Natural Language Processing (NLP) and Machine learning (ML) is particularly useful for the efficient and effective analysis of larger datasets, benefiting from the capabilities to identify patterns and themes. The techniques of NLP (Jones, Doane and Attbom 2021) and Topic Modelling (TM) (Jones 2021) were chosen to analyse the qualitative data in this research. Open Access code repositories provided a basis for developing specific algorithms to pre-process the data (data cleaning, formatting) (see references for **Software and GitHub repositories**) and perform the thematic analysis, which conformed to the principles of *Grounded Theory* (see, e.g., Charmaz 2006; Odacioglu

[64] https://help-nv11.qsrinternational.com/desktop/concepts/about_automated_insights.htm

and Zhang 2022)[65]. TM allows data to be statistically grouped and automatically labelled (tqx94 2022) in theme clusters based on keywords, which enable the automatic discovery of topics latent within or emerging from the data. Topic Modelling can provide insights and themes that were not anticipated in the initial design of this research (see results in **Chapters 4 to 6**) or might have been missed or overlooked[66]. Also, This information can be coded and formed within a thematic content analysis from the information provided by the communities – as some form of 'a lingua franca' (Mason 2002, 9).

The manual assessment or direct observation following from the TM, identify categories resulting from this research which will be correlated with current heritage categories as provided by the *Conservation Principles* of Historic England (English Heritage 2008) and HLC to assess a connection between historic landscape types and social values. This correlation is intended to assess the method's potential compatibility with current approaches and frameworks and address future demands on such tools.

The methods for map-based questionnaire design have also been adapted to achieve the aim of using freely available software packages[67]. Proprietary software, such as *Maptionnaire* developed by the Finnish company Mapita[68], has been used elsewhere (Dabaut 2021, 254). However, for this research Google Maps was integrated in *Qualtrics* survey, which provided the user interface for online data collection,

---

[65] Elements of Grounded Theory underpin the research to create categories based on the language used in the stories provided by participants, similar to the Alphabet of Distinctiveness (see **Chapter 1.2.2**).

[66] Justification for using the autocoding (topic modelling) in NVivo (see https://help-nv11.qsrinternational.com/desktop/concepts/about_automated_insights.htm

[67] Ideally, all software in this project would have been open source. However, for the survey, I had to resort to proprietary software (e.g., Google Maps, Qualtrics), which is free to use for projects like this, but not open source.

[68] It has been shown that participants using *Maptionnaire* needed guidance and skills to work with this interface, which led to lower-than-expected useful information and lower participation in the survey (see Dabaut 2021, 254).

allowing participants to set pins on a familiar map interface and add personal stories and information.

QGIS provides community-led software for the visualisation of spatial information. I used QGIS, adhering to the principle of using open-source software, to map and visually present the results of the three data sources. This research aims to produce maps for (1) heritage and landscape management and spatial planning purposes and (2) resources for engagement and participation of non-experts in landscape characterisation and identification of social values. Filtering for specific information, such as issues raised by residents of the PDNP and Sheffield, such information can be visualised as 'Issue Maps' highlighting the need for action to local authorities. Other examples include 'Hotspot Maps', which represent focal points for high footfall by visitors and tourists and potentially associated risks and opportunities.

The results of the analyses are provided in **Chapter 7.** Maps, representing people's values, needs and visions for the places they live, can be used by heritage professionals, for example, for planning, policy and management plan development, communication and outreach, and by local community groups to lobby for and better communicate their interests.

1.5.5 Sampling strategies

Sampling was influenced by the project partners and their support, e.g., providing opportunities to circulate the online surveys through their media channels and advising on potential partners for the in-depth interviews.

The social media sampling strategy allowed me to reach a wide range of people connected through posts related to the study areas. The decision to use Twitter (now X) as a data resource for social media analysis was influenced by the restriction of use and reuse and the internal structure in private or public groups of such datasets by other social media platforms, e.g., Facebook or Instagram. The opportunities for academic researchers, which provide access to historical data back to the beginning of Twitter in 2006 and the public posting nature of the platform, were granted through the Academic Developer Account. Access through this account allowed to collect Twitter data from three bank holiday weekends in the UK from 2019 to 2021. Initial inspiration for social media research was based on a workbook published as a GitHub repository by Bonacchi (reference in **Software and GitHub repositories**, Bonacchi 2021). This workbook provided a starting point for further development of social media research, particularly for adopting and adapting other methods, e.g., an emoji-based sentiment analysis (reference in **Software and GitHub repositories,** also Omkar 2019, Hutto and Gilbert 2020). The sampling of social media data has the disadvantage that it is biased through the focus on a particular group – Twitter users – and the nature of posts – showing life in a more positive way. However, the advantage of the method was the opportunity to gain a broad insight into the phenomenon and the large data availability.

Survey sampling was based on the principle of a *Convenience Sample* – a non-probability sample method (see, for example, Golzar and Tajik 2022). This sampling strategy was chosen to gain an insight that reflects the typical database of the project partners and benefit from the established connections of the partner organisations – the Peak District National Park Authority (PDNPA) and the South Yorkshire Archaeology Service (SYAS) of the city of Sheffield. The partnership allowed me to circulate the online questionnaire via the social media channels (Facebook and Twitter) of the PDNPA and an existing mailing list of the city of Sheffield, which consisted of over 4,000 residents who had agreed to participate in surveys. While this sampling strategy has drawbacks, e.g., sample bias and being less representative than other methods, it has significant advantages, e.g., time- and cost-effectiveness and ready accessibility. This form of sampling was expected to provide a larger database, required for the landscape-scale approach taken in this research, than would have been possible through other sampling strategies.

The interview dataset was compiled based on a *Typical Case* sampling strategy (Creswell 2017, 159). The PDNPA suggested potential participants based on their knowledge of and particular association with the area. The aim was to gain deep insights into the connection between people working and/or living in the PDNA and their personal stories of attachment with and relationship to the study area. This sampling strategy allowed me to interview a wide range of people from different walks of life, age groups, and experiences in the study area.

1.5.6 Ethical issues and procedures

The methods applied in this research were subject to rigorous ethical assessments by the University of York (for the ethical application see **Appendix 6**). As the methods developed during the research, including additional data sources, such as social media data, and techniques, such as using AI tools, the ethical approval was amended accordingly over time. The subsequent additions were assessed and approved (see **Appendix 7**).

The use of data from living people requires due diligence, data security procedures, and safe data handling (the required documents, e.g., consent forms and information sheets for participants of the survey and interviews, can be found in **Appendices 3 and 4**). Where possible, anonymity was ensured by using and publishing only synthesised data. For example, quotes from social media posts were not published, as their origin can be easily identified through an internet search. Similarly, only survey information was used that would not allow the identification of individual participants. In contrast, the interviewees waived their right to anonymity as the information given was presenting a deep insight into the individual and personal connection between the participants and their favourite places, which would have made anonymisation difficult, if not impossible.

Ethical issues arose, in particular, from the use of AI tools (for a detailed discussion, see **Chapter 7**). The techniques used in this research were based on both supervised and unsupervised learning methods. For example, Topic Modelling was based on a statistical method with no introduced bias through pre-labelled training data (**see Chapter 4**). In contrast, the sentiment analysis tool used in the social media research introduced this bias through annotated training data (**see Chapter 3**).

Further ethical implications were encountered from the potential use of models created based on the survey data. **Chapters 7** and **8** detail the concerns about decision-making based on AI tools' models and predictions and the introduced bias and limitations of the methods. **Appendices for Chapter 8/Appendix 7** also provides an addition to the ethical approval, outlining the solution for this issue.

The 'human-in-the-loop' principle was followed to ensure ethical integrity and assess the performance of the AI tools in this research. This was why datasets were small compared to those commonly used in AI approaches. AI is particularly useful for analysing large datasets that would take extensive time to annotated and assess. However, the clearly limited dataset used in this research allowed a thorough familiarisation with the data and processes during the analysis. For example, results of the AI analysis were compared with the manual analysis, and discrepancies were investigated (see **Chapter 3**, Sentiment analysis based on emojis as compared to text).

The following section will provide an overview of the structure of this thesis, which is based on a series of peer-reviewed and published papers resulting from this research.

1.6 Outline of the thesis

This thesis comprises a series of papers published as part of this research. **Chapter 2** will provide a wider background and literature review on specific elements of this research. The sequence of the papers in **Chapters 3** to **7** reflects the steps and sequence in which this research was conducted. **Chapter 3** will present the first paper focussing on one of the data sources – social media data. This chapter will also provide interim results and outputs of the analysis and discuss potential applications of the method. The method is based on information about favourite places and associated sentiments from a wide spectrum of social media users. In this chapter, I present Natural Language Processing (NLP) for data analysis. The following chapters will build on and extend this technique. The chapter also introduces various forms of visualisation useful for planning, management, and public engagement. **Chapter 4** narrows the focus of participants to people living and working in the study areas and introduces the data collection method of questionnaires in the form of online surveys. The chapter will introduce Topic Modelling (TM) for data analysis based on elements of *Grounded Theory*, which allows for the exploration of qualitative data that is free from preconceptions and predefined codes. This paper focusses on the thematisation and categorisation of reasons behind place attachment based on individual/social values in everyday living and working landscapes related to heritage and the historic environment. The chapter presents the survey data in a format similar to the social media research in **Chapter 3** and expands the visual capacities to reflect the deeper information basis presented by this data source. **Chapter 5** further narrows the spectrum of participants while, at the same time, deepening the exploration of reasons behind a strong place attachment and development of social values by using the method of in-depth interviews. Using the methodology deployed in the previous papers – NLP and TM – this represents an application of the techniques to interview data. The focus of this chapter is to provide a method for the thematisation and categorisation of reasons for rootedness and strong connections to place. With the focus on the categorisation concept, this paper applies a similar technique to a different data source as used in **Chapter 4** and draws from the previous analysis experience. **Chapter 6** brings the various data sources together and interprets what knowledge and understanding can be drawn from the analysis of social media, surveys, and interviews while, at the same time, putting the results of the study in the wider context of current and future applications. While the limitations and biases specific to each data source and method are provided in the respective chapter, **Chapter 7** presents underlying ethical implications and limitations associated with AI technologies and tools as a critical reflection of the research conducted based on these techniques. While AI has been deployed for the analysis of image-based data in archaeology for decades, the deployment of AI tools on text-based data in archaeology and heritage management is a recent development, at the time of writing. Considerations regarding ethical implications, which posed new challenges and risks for research and application, were a strong factor in this research – on the one hand, it provided new opportunities for data analysis and interpretation, and, on the other hand, it brought new challenges for data safety and transparency, inclusivity and explainability of the processes in view of fair and risk-aware use of the outputs of such analyses. The chapter identifies and describes the wider repercussions of automation and computerisation of processes in the discipline and beyond.

How the results presented in the publications have met the aims and objectives of the research project will be discussed in **Chapter 8**. The focus of this research, which partnered with the Peak District National Park Authority, UK, and the South Yorkshire Archaeological Services of the city of Sheffield, UK, lies firmly on the potential for practical application in a real-world scenario. Furthermore, **Chapter 8** will summarise the key achievements of the study and provide an outlook for further work in the field of social values and AI applications in archaeology and Cultural Heritage Management. While further developed techniques, such as deep learning and neural networks, including Large Language Models, lay beyond the scope of this research, I provided further opportunities for developing this method to an automated categorisation of qualitative data based on trained models and supervised learning.

The **References** chapter will provide the bibliography for all chapters, with a separate section for **Software and GitHub repositories**. The **Appendix** will provide supplementary material on the papers presented in this thesis and the wider research. Furthermore, it will give additional information and code not included in the published journal articles in the form of a lab book.

Chapter 2:

Literature review

2.1 Introduction

This section will position this research within the context of the historical and current heritage discourse[1]. First, a focus on social values will introduce national and international agendas and charters concerning the integration and development of the social value and everyday heritage in the heritage sector. Social values will be defined and the dilemma of categorisation elaborated (Fredheim and Khalaf 2016, 468; Mason 2002, 9; Rudolff 2006, 2; Stephenson 2008). Also, the issue around the use of jargon and the language of distinctiveness, or 'narratives of identity' (Rudolff 2006, 229) within historic environment datasets will be discussed (see, for example, Common Ground 2006; Fairclough et al. 1999, 12; Grove-White 1996, 9).

Then, the concept of place and place attachment will be defined and the topic of landscape-scale mapping applications will be discussed in view of people-centred, place-based approaches, visualisation of a 'sense of place', and character or distinctiveness. Tools such as Historic Landscape Characterisation (HLC) and Landscape Character Assessment (LCA) will be introduced as examples of landscape assessment and interrogation. In the section on socially sustainable development and planning, visualising narratives to create social value-based maps, will show a way toward the integration of local expertise in 'officially authorised' heritage assessment frameworks, combining expert views with local knowledge.

This chapter will give the necessary background for the following papers, which touch on the more detailed descriptions of landscape and heritage management concepts as elements of this research.

2.2 Charters and agendas – towards social value

2.2.1 Heritage and inherent value: The old model

Heritage assessment and management rely on the categorisation of values defined in a

---

[1] See also **Chapter 6**: Social Landscape Characterisation: A People-Centred, Place-Based Approach to Inclusive and Transparent Heritage and Landscape Management

canon of values ascribed to monuments, buildings, places, and landscapes that have, in one form or another, meaning and potential to contribute to the 'cultural capital' of society (Byrne 2008, 159; Harrison 2010, 243). The definition of heritage and the value system were developed from the first antiquarians' work in the 17th and 18th centuries, with a systematic structuring of heritage in typologies and seriation in the 19th and 20th centuries (West 2010, 9). This was the environment in which William Morris and John Ruskin became influential in establishing the historic and aesthetic value inherent in the 'authentic' fabric of buildings in the second half of the 19th century (Jones and Leech 2015, 7; Smith 2006, 89-90). Ruskin also had a vital interest in 'educating the people' – a notion that re-emerged in the 20th century, promoted by English Heritage and continued by Historic England (Byrne 2008, 168).

As part of the Ancient Monuments Protection Act 1882, a schedule introduced the notion of preserving monuments as islands of importance cut out of the wider landscape. This act and legislation of designation that followed started a process of valorisation and significance assessment with an emphasis on monuments of 'national importance'. However, this concept also created landscapes of insignificance and unrecognised 'unofficial' aspects of the historic environment (Harrison 2010, 240; Ireland, Brown and Schofield 2020).

The devastation of the two World Wars led to the establishment of the first charters concerning the historic environment with an initial focus on buildings and structures. The Athens Charter for the Restoration of Ancient Monuments of 1931 (ICOMOS 2021b), as well as the Venice Charter of 1964 (ICOMOS 1964) and UNESCO's Convention Concerning the Protection of the World Cultural and Natural Heritage in 1972 (UNESCO 2021a) continued the concept of 'authenticity', relying on the historic, scientific, and aesthetic value inherent in the material structure of ancient buildings, monuments, and sites (Ahmer 2020, 151; Jones and Leech 2015, 7-8; Jones 2017, 23).

2.2.2 Heritage and social values: The new model

The 'cultural turn' in the second half of the 20th century changed the understanding of heritage values[2] (Bonnell and Hunt 1999; Cosgrove 2004; for political and economic consequences, see Torre de la 2002, 3; Pearson 1995, 126; Pendlebury and Gibson 2016, 1-2). The emphasis shifted to a 'local' theory instead of a 'grand' theory, meaning instead of objectivity, identity and materiality, the focus was on hybridity and relations (Cox 2014, 105).

The second half of the 20th century was dominated by the development of charters and agendas to integrate 'everyday heritage' and the perceptions of lay people in heritage and landscape management. The conventions of the *Council of Europe* (CoE)[3] became influential in the reorientation of value definitions[4]. Furthermore, the Amsterdam Charter of 1975 (European Charter for the Architectural Heritage) introduced key principles that would become the foundations of heritage thinking in the decades to come by introducing the principles of:

- 'heritage as capital of irreplaceable spiritual, cultural, social and economic value' (Council of Europe 1975, Article 3)
- people having 'an instinctive feeling for the value of this heritage' (Council of Europe 1975, Article 2)
- the importance of heritage to contemporary society, the broadening of the view from important monuments to their wider setting (Council of Europe 1975, Article 1)
- the contributions heritage can make to a 'harmonious social balance' (Council of Europe 1975, Article 4)

Additionally, the Dresden Declaration of 1982 introduced the principle that monuments can have 'symbolic value' and 'spiritual value' (ICOMOS 2021a, Article 7

[2] For changing understanding of values and the role of professionals as opposed to public participation, see **Chapter 4.2.**

[3] https://www.coe.int/en/web/portal/home

[4] For more on the CoE see **Chapter 4.2** and **B. Appendices for Chapter 4: Supplementary material 1**.

and 9)[5]. Also, the term 'place' was defined more broadly in the explanatory notes of the Burra Charter and included 'travel routes, a community meeting place, a site with spiritual or religious connection'. It acknowledges 'natural elements' or 'spaces and view that might be part of the significance of a place' (ICOMOS 2013, Explanatory Notes to Article 1). Regarding the participation of non-experts, the charter states:

> Groups and individuals with **associations** with a place as well as those involved in its management should be provided with opportunities to contribute to and participate in understanding the cultural **significance** of the place. Where appropriate they should also have opportunities to participate in its **conservation** and management (ICOMOS 2013, Article 26, Paragraph 3, original emphasis).

Despite weaknesses and room for interpretation[6], many experts have seen the Burra Charter as the adoption of social values into the canon of heritage values. The charter's influence on conservation practice and guidance in the UK was acknowledged as a 'watershed' moment encouraging the participation of everybody (Chitty 2016, 1). Particularly crucial in this quote is the mention of individuals as having an interest. As shown in **Chapter 6** and emphasised by Johnston (2023), the individual is not recognised in current approaches, which focus on project work with predefined groups and communities.

---

[5] For a critique of adherence to the principles of the Authorised Heritage Discourse (AHD) inherent in these documents, such as the Athens and Venice Charters mentioned before, as well as the Burra Charter (Australia ICOMOS Charter for Places of Cultural Significance, first adopted in 1979 (ICOMOS 1979); Australia ICOMOS and International Council on Monuments and Sites 2013 (ICOMOS 2013)) see Smith 2006, 113. These documents also strengthen the role of heritage experts (Smith 2006, 104).

[6] Important in this respect are, however, the parts in the paragraph that are not emphasised, and Smith argues that they reflect the key issues in line with the AHD in that the parties are supposed to be invited to 'understand' the significance rather than express their own experiences and views (Smith 2006, 104)[6]. Furthermore, the role of experts is again reinforced by the demand within the charter that 'supervision [...] should be implemented by people with appropriate knowledge and skills' (Smith 2006, 105)

Furthermore, in the US, the Getty Conservation Institute in Los Angeles undertook a research project into the value of heritage in the years between 1998 and 2005 (Avrami, Mason and de la Torre 2000; Torre de la 2002; Mason 1999). This project demonstrated a continuation of the value discourse and the importance of the historic environment in creating community values as an expression of a societal shift.

International conventions strongly influenced the UK's heritage practices and policies, introducing social aspects of everyday landscapes, local knowledge, and public participation. Both the European Landscape Convention (Florence Convention) (Council of Europe 2000) and the Framework Convention on the Value of Cultural Heritage for Society (Faro Convention) (Council of Europe 2005; see also Schofield 2015) stressed the importance of communities and individuals' rights to voice their views democratically and engage with heritage alongside the professional community. The latter also stressed the social value aspect more strongly but was not ratified by the UK. Jones and Leech (2015, 10) argue that these ideas still strongly influenced the English Heritage (and later Historic England) guidelines and advice in, for example, English Heritage's *Power of Place: A Force for Our Future* (English Heritage 2000), *Knowing your place* (English Heritage 2011a) and the *Conservation Principles* (English Heritage 2008)[7]. These documents focussed on the connection between people and places and active participation of communities in development of neighbourhoods. The Florence Convention was ratified and enforced in the UK in 2007 (Herring 2009, 68)[8] and emphasises, for example, the aspect of the landscape as 'everyday areas':

> Acknowledging that the landscape is an important part of the quality of life for people everywhere: in urban areas and in the countryside, in degraded areas as well as in areas of high quality, in areas recognised as being of outstanding beauty as well as everyday areas (Council of Europe 2000, Preamble).

---

[7] For more on the Conservation Principles see **Chapters 4 and 6.**

[8] HLC developed in the 1990s in the UK directly influenced the convention. Some of the people involved in defining the convention's objectives and recommendations also worked on characterisation projects, such as Fairclough (England) and Fojut (Scotland).

The beginning of the 21st century saw the erosion of 'hard lines' dividing the cultural and natural environment and the tangible and intangible, leading to the acknowledgement that values are created and negotiated in a dynamic process between people, communities, and places (Byrne 2008, 158; Jones 2017, 21; Turner 2018, 38-39). These strong dichotomies were questioned and subsequently reversed (see, e.g., Byrne and Ween 2015; Harrison 2015; Fredengren 2015). The dissolution of dichotomies in heritage thinking and categorisation influenced the discussion on categorisation per se (see **Chapter 2.4** for the categorisation dilemma of social values). The Faro Convention ended these strong dichotomies (Jokilehto 2016, 20) and advocated a holistic approach to cultural heritage by stating:

> Cultural heritage is a group of resources inherited from the past which people identify, independently of ownership, as a reflection and expression of their constantly evolving values, beliefs, knowledge, and traditions. It includes all aspects of the environment resulting from the interaction between people and places through time. (Council of Europe 2005, Article 2a).

2.2.3 Conclusion

In summary, the development of heritage values and categorisation has shown a trend from an expert-led, significance-based, site-oriented understanding of heritage and the wider historic landscape to community-focussed everyday heritage concepts. Social values were defined and advised in international charters and agendas and national policies and guidelines. People's visions and needs gained importance in a society that prioritised the social benefit of heritage and a sense of place, belonging and identity as a new benchmark for the preservation and conservation of the historic environment, including mundane and everyday things that people value in working and living environments. Historic England's Corporate Plan priorities emphasise the importance of 'connected communities' and 'active participation' (Historic England 2023). The topic of 'Everyday Heritage' has also been further developed at Historic England with the launch of several Everyday Heritage projects with a particular focus on working-class histories[9].

A range of diverse projects have shown how social values can be identified and meaningfully integrated into heritage and landscape management. **Chapters 1.2** and **1.3**. discuss examples of such community engagement and participation approaches. The next section provides an overview of social value development in research and the challenge in practical applications.

2.3 Social values

2.3.1 Definition

As shown above, the 'cultural turn' within the heritage sector signified a new understanding of the relationship between people and the historic environment (Johnston 1992, 7; 2023; Jones 2017, 21-25; Rudolff 2006, 4; Waterton, Smith and Campbell 2006, 393). The emphasis on values shifted towards the sense of place, belonging and identity. The significance of the historic environment for contemporary communities and forms of memory and spiritual association as key principles of the new engagement with everyday heritage gained importance. Also, the views of laypeople regarding their environment were treated as an essential part of understanding the quality of places. The notion of a 'sense of place', especially, has been the subject of extensive discussions within the heritage sector (Byrne 2008, 155; Clifford and King 1996, 10; Harrison 2010, 243; Jones 2017, 21; Jones and Leech 2015; Johnston 1992, 10; 2017; 2023; Meinig 1979, 3; Schofield 2007, 111). Place can be understood as an abstract construct of practice and experience different from locality (Pink 2012, 3), while a sense of place is created in terms of a spatial entity, material form, or specific environment. The human geographer Yi-Fu Tuan introduced the term 'sense of place' or 'Topophilia'[10] (Tuan 1977, 1) in human geography, expressing

---

[9] https://historicengland.org.uk/campaigns/help-write-history/everyday-heritage-grants/

[10] W.H. Auden forged the term in his introduction to John Betjeman's book 'Slick but not streamlined' in 1947 when he wrote: 'Topophilia differs from the farmer's love of his home soil and the literature's fussy regional patriotism in that it is not possessive or

emotions of security, warmth, and familiarity associated with a place. Such emotional connection can create social value, expressed through, for example, oral history, local history and genealogy, festivals, everyday practices, graffiti, traditions, and memorial events (Jones 2017, 25). They are also spiritual and religious associations, symbols, and subjects of official narratives and folktales (ICOMOS 2013, Explanatory Note, Article 1; English Heritage 2008, 31-32; Jones 2017, 24-25; Jones and Leech 2015, 33; Tiller 2020, 281). Social value and the reasons behind it form the local knowledge sources within communities (Tiller 2020, 1-3).

Jones and Leech (2015, 5) argued that heritage is ‘produced through experience and practice’ and continuously created and negotiated dynamically and fluidly (see also Byrne 2008, 169; Jones 2017, 21). The production of values depends on the meaning communities ascribe to their environment, and, in doing so, they create heritage with local or everyday significance. This is a ‘cultural and social process’ (Smith 2006,2) that can be understood as ‘heritage work’ (Byrne 2008, 171; Jones 2017, 24-25; Smith 2006, 1). Smith (2006, 59-60) referred to this concept as ‘memory work’, a personal, collective, emotionally charged affection different from official history.

2.3.2 Practical approaches to social value assessment

The 1990s saw an increasing interest in ethnographic studies in place, perception, and phenomenological experience of the environment (Pink 2012, 37). Several approaches in academic research projects showed the benefits of participatory and people-centred approaches (Cinderby, Snell and Forrester 2008; Cinderby et al. 2012; Jones and Leech 2015; Jones 2017; Nardi 2014). However, integration into heritage practice and management is still viewed as problematic for different reasons (Emerick 2016, 65-66). For example, social values are ‘soft’ and ‘fuzzy’ data, and collecting such ‘slippery’ data can be challenging (Pearson 1995, 156); they might be indecipherable, not easily

---

limited to any one locality […] it has little in common with nature love. Wild or unhumanised nature holds no charms for the average topophile because it is lacking history […]’. (Betjeman 1947, 11) understood (English Heritage 2008, 32), and not recognised by ‘outsiders’, especially heritage experts who ‘parachute’ in to assess the significance of an object (Emerick 2016, 75; see also Byrne, Brayshaw and Ireland 2003, 3). The reluctance to change and adapt practice has been seen as a factor in the slow integration of social values into the official practice (Emerick 2016, 65).

Public involvement and participation have become a focus for development over the last three decades, and the role of experts has been at the centre of discussions. As Smith (2006, 94) points out, the way heritage is managed influences public opinion (Avrami 2009, 179; Hølleland and Skrede 2019, 833; Schofield 2014; Emerick 2016; Jokilehto 2016, 31)[11]. The change towards inclusion and participation led to research and community projects with the creation of social value toolkits and community initiatives[12].

However, due to their subjective, qualitative nature, social values are inherently difficult to collect, analyse, and use in a system dominated by (apparently) robust, objective, measurable data sets commonly associated with heritage practice (English Heritage 2008, 36; Jones and Leech 2015, 15; Mason 2002, 9)[13]. Social values are now accepted in the heritage sector but remain a complex concept (Pearson 1995, 21). Weighing them against other ‘traditional’ values still marginalises this category (Byrne 2008, 156; Dalglish 2018, 20-21; Jones 2017, 25; Mason 2002, 10). The following section will introduce categorisation and the narrative approach as crucial elements for a systematic approach to social value integration into existing official assessment frameworks.

[11] For a discussion on the role of experts and laypeople in heritage assessment and decision-making see **Chapter 6.**

[12] For a detailed discussion of past and current projects and initiatives on social value assessment see **Chapter 1.2**.

[13] For the challenges of collecting social values, see **Chapter 3, 4,** and **5**.

2.4 Categorisation and narratives

2.4.1 Value categorisation – Background

Categorisation ensures a degree of consistency and reproducibility in assessment frameworks. From the beginning of the category systems in the 19th and 20th centuries, the definition of value categories aimed to make the historic environment comparable, decision processes consistent and replicable, and significance and conservation aims descriptive (Mason 2002, 9)[14]. However, the constraints of a rigid system applied to fluid and dynamic values have been found unsuitable, and other approaches or extensions to existing categorisations frameworks have been proposed. The categorisation of heritage values constitutes a considerable challenge within the heritage management sector. The balance between generalisation and particularisation, resulting in a too broad or too fine-grained category system, defies the notion of fluid and dynamic value creation within the historic environment.

After a period of increasing separation into broader categories, such as natural and cultural landscapes (West 2010, 3; UNESCO 1972), or tangible and intangible heritage (UNESCO 2018), such a division's nonexistence in the real world has been widely acknowledged. Subsequently, the trend shifted towards a holistic approach to heritage (Byrne, Brayshaw and Ireland 2003, 24; English Heritage 1997, 3; Harrison

---

[14] Alois Riegl, an art historian and general conservator in Austria, established the first systematic categorisation of heritage values, creating a framework for heritage professionals (Ahmer 2020, 151). Riegl defined two categories of value at the beginning of the 20th century which became canonical for the heritage sector: memorial values (age, historical and intended value) and present-day values (use, art, newness, and relative art values) (Ahmer 2020, 150; Pendlebury and Gibson 2016, 6-7; Riegl 1903; Walter 2014, 634). Historically, value that would determine preservation was defined by an elite group of professionals who 'relied on ideas of selection and classification eventually expressed in state-defined and controlled lists, and on principles of conservation' (Pendlebury and Gibson 2016, 6-7).

2010, 2015; Pearson 1995, 315; Rudolff 2006, 2). Also, places often have an intangible aspect, and vice versa, intangible values can have a spatial connection (Kaufman 2013, 20). Therefore, new valorisation models were needed to include these new ideas and aspects of heritage and landscapes, which will be detailed in the next section.

2.4.2 Development of new categorisation approaches

Fredheim and Khalaf (2016, 468, Table 1) produced a historical overview of the various 'typologies for cultural heritage'. The comparison shows that this systematic organisation of values became increasingly inflexible over time, leading to a canonical, traditional, conservative view on the historic environment, and thereby limiting social value's flexibility and dynamic character. The traditional system proved outdated and not suitable for the postmodern view, leading to a rethinking of the value categories, the methods to capture these, and the question of whether categorisation itself is useful.

Several value systems have been developed (Fredheim and Khalaf 2016, 468; Mason 2002, 9) as a basis for significance assessments. These are firmly rooted within the traditional thinking developed in the 18th and early 19th centuries while adapting to the needs of the time. However, as Smith (2006, 105, 299) has pointed out, real change within the profession has not prevailed over the canonical categories and the weighing of values against each other. Stephenson (2008) suggested a *Cultural Value Model* that breaks up traditional value categories. Stephenson's model breaks with the traditional categories by dissolving the value categories of historical, evidential, aesthetic, and communal values and reorganising them based on 'relationships, forms and practices' (Stephenson 2008, 134 -135, and 134, Fig. 2). Stephenson not only reorganises existing value categories, but also adds new concepts. For example, he includes categories of natural landscape aspects, social or community aspects as traditions and practices, and relationships as memories, senses of place and attachment[15].

[15] For the extension of categories in a flexible system, compare **Chapter 4.4.2**.

This model offers less rigid compartmentalisation of the different aspects, organising values by dissolving the traditional value categories and opening them up to a more holistic view of the landscape. It also reinterprets time-depth by differentiating surface values of the present from embedded values of past cycles of value creation (Stephenson 2008, 135-136, and 137, Fig. 6). The model offers a more open, flexible, and inclusive approach to value assessment and includes additional intangible aspects of the landscape that are often elusive and difficult to integrate into the practical process of landscape assessment and management (Stephenson 2008, 136). This approach is supported by the notion of Clifford and King (1996, 9), arguing that only outside of the 'official' frameworks are people permitted 'to devise **their own** and shared categories of meaning' (emphasis by the author), as they were at that time. Jones (2017, 22) also questioned the traditional approach as it 'tends to objectify and fix different categories of value'. Schofield (2007, 113) emphasised the spatial dimension of meaning as 'activities conducted in particular places (points), generic activity at a landscape scale (space or areas), and activities that are not place-specific and relate more to temporality (trajectories or the lines of journeys)'. Such an approach emphasises the spatial connection between the abstract concepts of meaning, time, and place.

Most of the approaches presented in this overview attempt to introduce a level of flexibility to the field of social and heritage values but still predefine value categories[16]. While the approaches are expert-led and top-down, heritage values must be categorised in a bottom-up approach to truly reflect the value heritage and the historic environment have for the people. Reflecting people's needs, visions and aspirations should be grounded in and emerge from people's personal connections to places, using language that stems from such experiences and attachments. A method that allows capturing of latent themes or dominant topics in the narrative of individual

[16] Compare the Cultural Mapping project for the identification of tangible and intangible aspects of places (https://bangkok.unesco.org/content/cultural-mapping) (Currie and Miranda Correa 2021; Currie and Correa 2022; McKeithen 2015).

experiences, conveying the reasons for valuing or attaching to places, would benefit from a *Grounded Theory* approach (Charmaz 2006; Odacioglu and Zhang 2022)[17].

The next section will elaborate on the challenge of introducing a common language that can benefit communication between laypeople and experts.

2.4.3 Narratives and new language

The role of the heritage professional and public engagement methods have changed over the past decades. An important aspect of engaging the public is adapting the language to make heritage more accessible for lay people. Local heritage and landscape management need a common language to negotiate and express values and aspects of heritage (Byrne 2008, 165). Commonly heritage practice is dominated by jargon that excludes terms of social values and non-expert opinions as expressed by laypeople. Unfamiliar jargon used by experts in assessing the historic environment makes it more difficult for laypeople to understand and engage with heritage (Grove-White 1996, 9). A new approach to heritage practice can help dissolve the strict division of experts and laypeople and encourage dialogue and cooperation (Schofield 2015b; 2016) based on the narratives and stories of ordinary people in their everyday lives and lived-in worlds (as discussed in **Chapter 1.5.4**). For example, the 'Alphabet of Distinctiveness' articulated of essential terms of the meaning of places and formed the basis for creative writing about local places (Common Ground 2006; see also Crouch 1996, 22; Hayden 1995, 64). Using it as a basis for a terminology closer to a common language that would be understood and making datasets, such as Historic Landscape Characterisation, more accessible for laypeople, was suggested (Fairclough et al. 1999, 12). A new alphabet of local distinctiveness might help create a more inclusive language as it is based on the stories of the people (see Fairclough et al. 1999, 12; Hayden 1995, 66). This approach is similar to what Rudolff (2006, 229) referred to as 'narratives of identity' from which a value system can be created.

[17] For a discussion on Grounded Theory and the use of emerging themes from empirical research, see **Chapters 4** to **6**.

A different language can also mean other ways to express the meaning of landscapes, and art and performance have been proven as media that can share such intangible ideas of places and landscapes with 'outsiders'. Such expressions can encompass art installations such as the Chiswell earthworks (Hayden 1995, 64-66; Schofield 2007, 110), drawing and mapping as in the Parish Map projects (Common Ground 1996), creative writing, photography and videos, and performing arts (Pearson and Shanks 2001; Smith 2006, 66-67). In a similar way, Natural England suggested including information drawn from music, literature, and art to capture a different side of the landscape perception from a non-expert heritage professional position (Tudor 2014, 50), a concept that Common Ground's Parish Map project has realised since the 1980s (Perkins 2007, 130).

Such projects offer a creative way of expressing 'soft' values that give an insight into the world of locally held values that are usually not visible to heritage professionals or outsiders (Harrison 2010, 261). Collecting, analysing, meaningfully integrating and practically applying this data requires thinking outside the box, cooperating with non-experts, and applying interdisciplinary methodologies (Byrne 2008b, 150-151). Rudolff (2006) argued the need for heritage professionals for 'active listening, trans-cultural empathy, and ability to meaningfully re-narrate expressions of identity' (Rudolff 2006, 233). However, the question is the degree to which these artistic expressions of distinctiveness can be meaningfully integrated into local planning and development decisions[18]. It would require practitioners to be trained and open to new practice methods. Current projects show a trend toward developing strategies to include such new approaches in planning (see **Chapter 1.2**).

2.5 Place attachment and social values

2.5.1 The concept of place

Social values are closely related to places and the memories, experiences, and

[18] For a discussion of projects, including participatory and artistic approaches, see **Chapters 1.2 and 6.4**.

perceptions people develop in connection with them. Understanding place attachment will, therefore, help to gain a deeper understanding of how social values develop with place attachment. This section will introduce the research on place attachment and methods and theories that relate to landscape-scale attachment, preference and, perception of people within the historic environment.

Intensive research over the past 50 years, explored the roots and meaning of the connection between people and place (Brown and Raymond 2007; Brown, Raymond and Corcoran 2015; Lewicka 2011b, 2011a; Maguire 2017; Manzo and Devine-Wright 2014; Manzo and Perkins 2006; Proshansky, Fabian, and Kaminoff 1983; Raymond 2013; Rollero and De Piccoli 2010; Williams 2000; Williams and Vaske 2003). Reviews on the people-place-relationship or place attachment theory, compiled by Altman and Low (1992) and Lewicka (2011b), provide a comprehensive summary of research done since the emergence of place attachment research as a subject of interest in the 1970s. While Altman and Low reflect on the various opportunities to interpret place, landscape, and environment, Lewicka elaborates on methodology and the theory development.

The concept of place is not yet fully understood (Maguire 2017, 5) despite the attempt to define place by various disciplines, for example, Geography (Cresswell 2015; Tuan 1980; Relph 1976, 2021), Urban Planning (Hayden 1995; Lynch 1960), Environmental Psychology (Scannell and Gifford 2010), Landscape Architecture (Thwaites 2001; Thwaites and Simkins 2006), Anthropology (Ingold 1993, 154-156), and Archaeology (Tilley 1994, 14-20). Place can be defined from the perspective from which it is viewed: (1) the place as physical reality with a quality or 'genius loci', an environment that has affordance for a connection – an agency of the place, and through natural topography, trees, rivers, views or tranquillity (Kaplan 1984; Stedman 2003; Stepanchuk, Gafurova, and Latypova 2020; Wuisang 2014); (2) the product of people's perception – a social construct (Byrne 2008, 154; Milligan 1998). To understand this complexity, it is important to acknowledge the mutual influence and agency of the natural environment and people on each other – the physical and social attributes of place (Proshansky, Fabian, and Kaminoff 1983, 64). Heritage plays a central role in the process of creating a sense of place and 'is both an input and an output of the process of heritage creation' (Graham, Ashworth, and Turnbridge 2000, 4). In this thesis, 'place' plays an essential role as a general term for the various elements of the historic environment, such as buildings, sites, monuments, objects, landscapes, and natural areas.

One concept of place is defined as 'genius loci', which interprets place beyond political and social constructions as an inherent quality or essence; it acts like an active agent in the creation of a sense of belonging and identity (Wuisang 2014). Stedman (2003) demonstrates that landscape elements and environmental attributes are essential for constructing meanings and not exclusively social. Relph (1976, 31) emphasises the importance of natural affordances of places, particularly views, as factor in place preference and the creation of attachment.

Therefore, to understand place attachment in theory and practice – to use place attachment effectively and meaningfully for planning and development – two aspects of place should be considered: the physical location of the place – including the natural and environmental landscape aspects – and the emotional attributes of people, which create meaning through individual connection to places.

Place, in terms of spatial and physical expression, can be defined on various scales: as a building, street, neighbourhood, region, landscape, county, or country (Shamai 1991) – even extending to 'a pale, blue dot' (Sagan 1990) that is threatened by climate change with huge implications for the residents of the planet – and people's connection to these places can have various reasons. Some places can be pinpointed on a map with an exact geographical position; others might be defined as areas with 'fuzzy' edges (Maguire 2017).

2.5.2 Sense of place and social values

People connect to places for different reasons and develop specific forms of sense of place (Feld and Basso 1996), belonging and identity. A place acquires meaning for people when it affords tangible and intangible benefits. These benefits may be security and familiarity, rootedness and belonging, shelter and food, work and community participation, positive feelings, and space for mental and physical health. Social value is created when the meaning of the qualities of a place are being weighed and signified

(Williams and Patterson 1999, 142). Prohansky, Fabian, and Kaminoff (1983) emphasise the importance of perceptions and experiences of a person to develop place identity. Stefaniak, Bilewicz, and Lewicka (2017) showed the development of a place identity as the relationship between people's personal life stories and the history and past of a place based on ancestry, memories, and traditions.

The reasons people connect to places have been subdivided into the concepts of place dependency and place identity (Williams and Vaske 2003, 831). While attachment is seen as connections to places based on aesthetic and recreational benefits (Williams 2000), place dependency, similar to the territorial behaviour of animals, emphasises the utilitarian aspect of place, such as work, food, and education (Brown, Raymond and Corcoran 2015). However, both elements contribute to the quality of places and fulfil both purposes. For example, a community market can function as a place to purchase food and provide a community meeting place that is imbued with meaning and social value.

Place attachment and place identity play a vital role in quality of life, place satisfaction, and overall emotional connection to a place, which influence a resident's environmental behaviour (Scannell and Gifford 2010), trust in local authorities and development of an identity based on the history and past of the place (Stefaniak et al. 2017). Place attachment has been shown to be independent of social status and economic circumstances. For example, Shamai (1991, 356) asserted that 'homeless' people in modern countries do not develop any sense of place; this hypothesis has since been debunked by Kiddey (2014) and her research with homeless people, showing their strong connection to specific places in Bristol and York.

In relation to historic places, English Heritage (2000, 2) and the National Trust (2017) undertook surveys to measure the influence of heritage on the feeling and connection of visitors to historic sites and buildings. These examples proved the significance of archaeology and heritage in regard to the development of place identity and identity based on cultural belonging and not 'legal membership' (Hayden 1995, 8). The European Landscape Convention (Council of Europe 2000, Article 5) states:

> ‘to recognise landscapes in law as an essential component of people’s surroundings, an expression of the diversity of their shared cultural and natural heritage, and a foundation of their identity’.

Like in the case of communal markets, traditions, memorials, and festivals, such as ‘Well Dressing’, ‘Sword Dancing’ and the ‘Castleton Garland Festival’, can function as important actions to reinforce the bonds of the community and the connection to the place (Byrne 2008, 167)[19].

Detailed studies of the reasons behind place attachment may allow identification of the important question of what makes places distinctive and reveal predictors for specific behaviour or benefits. Such particular studies may potentially reveal the deeper meaning of places that are often unconscious even to the people themselves (Williams and Patterson 1999, 153). However, in order to understand place attachment as a social value, the people-place connection should be approached holistically for a better understanding of the phenomenon. In this sense, Lewicka (2011, 208) argues that the concepts of the connection between people and places, such as sense of place, place identity, rootedness and place satisfaction, should be reconnected to gain the full view of the problem. Therefore, it is necessary to acknowledge that place attachment creates bonds between people and places that are often not recognised by local authorities, planners, and heritage managers. Change and development can disrupt this connection and lead to the feeling of being cut off from familiar and secure structures (Proshansky, Fabian and Kaminoff 1983, 67), and a background of social value within communities may have the potential to provide socially sustainable development.

Chapters 5 and 6 will discuss place attachment in relation to this research, putting the method developed and used in this project and the results into the context of the wider research environment. The next section introduces the concept of landscape and characterisation, which provides context for the approach taken in this

---

[19] See also **Chapter 5**.

research, particularly in view of the scale on which social values can be collected and represented.

2.6 Landscape, maps, and characterisation

2.6.1 The concept of landscape

Landscape[20] is a term whose definition has changed over time and varies across countries (Turner 2007, 40). Tuan (1979, 90) defines *landscape* as 'ordering of reality from different angles'.

The human geographer Carl O. Sauer was influential in the concept of cultural landscapes with his work on the 'Morphology of Landscapes' in 1925 and shaped the understanding of landscapes in the early 20th century (Sauer 1996). He rejected environmental determinism, which was the predominant paradigm of his time, and instead emphasised the importance of human agency following the German school of geography (Rogers, Castree and Kitchin 2013). Sauer defined landscape very much in the terms by which it is understood today, arguing that there is almost no natural landscape left in the world and describing landscape as successive layers of human activities (Sauer 1996, 307).

Furthermore, he defined *culture* as 'the impress of the works of man upon the area' (Sauer 1996, 303), emphasising the human agency shaping the environment. This idea of landscape was developed during the early 20th century with aerial

[20] The Oxford Dictionary of Human Geography defines *landscape* as: 'The arrangement of pattern of things on the land' and 'the terrain, shape, and structure of land' and also: 'the social and cultural significance and meaning of such patterns and terrains' (Rogers, Castree and Kitchin 2013). Furthermore, the Oxford Dictionary of Geography states that the landscape is 'compromised, partial, contested and only provisionally stable as modes of ordering the world and our engagement with it' and 'landscape as a social space' (Mayhew 2009).

photography and landscape investigation pioneered by OGS Crawford who coined the term 'palimpsest' in archaeology:

> 'The surface of England is a palimpsest, a document that has been written on and erased over and over again; and it is the business of the field archaeologist to decipher it' (Crawford 1953, 51).

This influential view of landscape is based on 'the fact that all present landscapes are the result of all their history' (Bowden 2001, 41). The contribution of landscape history developed by WG Hoskins in his influential book, *The Making of the English Landscape,* in the 20th century (Hoskins 1955) shed new light on the processes by which landscapes were formed. This idea was set out as a key principle of HLC in the 1990s.

In line with this view, the European Landscape Convention (ELC) defines landscape as:

> '... an area, as perceived by people, whose character is the result of the action and interaction of natural and/or human factors.' (Council of Europe 2000, Article 1a)

Landscape research and disciplines have since accepted the definition of landscape as a sum of all the parts that had been treated separately before. In this holistic approach, landscape encompasses the natural, cultural, and perceptual/aesthetic aspects (see Tudor 2014, 9 Fig. 1; also Brown, Mitchell and Beresford 2005, 4; Thwaites 2001, 254).

In summary, it is widely acknowledged that landscape is a social or cultural construct created through: (a) human activities in the past tangible in contemporary landscape; and (b) the perception or intangible aspect of the material form (Bradley et al. 2004, 6; Byrne, Brayshaw and Ireland 2003, 50; Byrne 2008, 155; Darvill 1999, 107; Fairclough 2002; Phillips 2005, 20; Schofield 2014, 2; Turner 2018, 39). It is, therefore, crucial to understand the 'symbolic creation of landscapes, the cultural meaning of aspects of the physical environment [...] and the values and beliefs that sustain these symbols and their meaning' (Greider and Garkovich 1994, 21; see also Rudolff 2006). As stated in the Oxford Dictionary, people see the world in patterns to understand it better (Jeffries 2012, 125) – a concept that is reflected in the maps of HLC (see **Chapter 1.2.6**).

2.6.2 Tools for landscape characterisation and assessment

Rural and urban landscapes consist of elements that contribute to distinctiveness and a sense of identity for the local community. Historic England and Natural England developed tools for characterisation and assessment of wider landscapes to aid sustainable planning and decision-making and facilitate change and development. Two such tools are Landscape Character Assessment (LCA) and Historic Landscape Characterisation (HLC) (and Historic Land-use Assessment (HLA), an equivalent developed by Historic Environment Scotland[21]). The following will introduce both tools focusing on HLC and compare and contrast the two methods.

LCA was developed to identify the variations or distinctiveness of landscapes based on their particular character (Tudor 2014). The first work in the field of character assessment started in the UK at the Countryside Commission in the 1980s, with the first guideline published in 1993[22]. The technique is usually applied as a reaction to planned development and allows a degree of community involvement. However, data consulted for the projects were mainly desk-based or expert-led and assessed.

HLC was developed by English Heritage (now Historic England) in the early 1990s to capture and visualise the character of wider landscapes in a generalised form and continuously represent the time-depth across the historical landscape, which was created by human activity and natural processes (Fairclough et al. 1999, 4-5; Turner 2018, 40). HLC includes rural and urban landscapes and expands to the 'ordinary, the degraded and the modern' (Herring 2009, 68).

---

[21] https://hlamap.org.uk/content/about-hla

[22] Countryside Commission (1993) Landscape Assessment Guidance, CCP 423, Countryside Commission, Cheltenham.

Key principles of HLC include the **general** and **value-free** representation of a **continuous** historic landscape as it exists in the **present** time (time of the respective projects) bearing **traces of the past**.

After decades of intense focus on conservation and significance assessments of sites, monuments, and buildings – effectively creating islands of preservation in the landscape – Historic England's approach extended, first, to the setting of the designated assets and, second, to the acknowledgement of the landscape as ubiquitous (Herring 2009, 68; Meinig and Jackson 1979, 2; Turner 2007, 40) and imbued with character and historic time-depth, leaving no blank or 'grey' areas (Fairclough et al. 1999, 1-2; Clark, Darlington, and Fairclough 2004, 6; Turner 201, 40). This concept was developed as a key element of the emerging discipline of Landscape Archaeology, starting in the 1970s and 80s (Fairclough et al. 1999, 4). The discipline championed the widely accepted position that a site- and monument-based approach does not adequately represent a sustainable approach in heritage management (Byrne, Brayshaw and Ireland 2003, ix; Darvill 1999, 106; Fairclough and Barnatt 1999, 7; Fairclough 2007, 86; Fairclough et al. 1999, 4). Therefore, HLC based the principle to get the 'big picture' on the concept underpinning Landscape Archaeology designed to work on large scales, rather than on a local basis, diverging from the outdated, narrow view of earlier understanding of monuments and settings (Countryside Agency 2002, 6; Fairclough et al. 1999, 7).

The key principle to present the historic landscapes in their current form, containing the traces and processes that formed them, illustrates the time-depth – the stories of past generations that changed and shaped the environment through their actions and decisions. Equally, today's actions and decisions will have an impact on the landscapes that will be inherited by future generations. Today's decisions will have consequences for how heritage and cultural landscapes will be preserved or adapted (Clark, Darlington, and Fairclough 2004, 6; Fairclough and Barnatt 1999, 6; Fairclough et al. 1999, 8; Turner 2018, 42). Turner (2007, 46) argues that landscapes are dynamic and ever-changing, and the historic environment's management has to adapt to this. Furthermore, HLC attempted to capture the distinctive character of landscapes as a reaction to the changing attitudes expressed by, for example, the ELC. The ELC conveyed it as a continuous definition of landscapes 'as perceived by people', including the elements of landscape that are not officially designated or recognised as heritage (Council of Europe 2000, Article 1a; see also Fairclough 2007, 84).

HLC was developed in a way closely related to LCA and adopted some of the methods. For instance, spatial units are represented as Historic Landscape Character Types categorising the varied character in the landscape (Turner 2018, 41) on a county-based scale (Fairclough 2001, 25). The character types were determined by interpreting processes that formed recurring patterns in the landscape, which could subsequently be generalised across regions (Clark, Darlington, and Fairclough 2004, 7; Fairclough 2002, 280). While this generalisation allowed the identification of patterns, distributions, similarity, and distinctiveness, such an over-simplified approach was critiqued by Williamson (2007) for its bird's-eye view. He argued that the detachment from the object – the physical landscape – made HLC maps incapable of representing the essence and subtle nuances of a landscape that create genuine distinctiveness, which is only possible from a ground perspective (Williamson 2007, 67, 69-70). Sauer had noted in this respect:

> 'An ordered presentation of the landscape is a formidable undertaking. Beginning with infinite diversity, salient and related features are selected in order to establish the character of the landscape and to place it in a system. Yet generic quality is non-existent in the sense of the biological world. Every landscape has individuality as well as relation to other landscapes, and the same is true of the forms that make it up' (Sauer 1996, 30-31).

Initially developed for rural environments, the idea of characterisation was later also applied to urban areas (Thomas 2006) and as Historic Seascape Characterisation (HSC) (Hooley 2014; Turner 2018, 45) and Seascape Character Assessment (SCA) (Tudor 2012) to maritime environments. A programme focusing on the characterisation of the late 20th-century landscape was the *Change and Creation* programme led by English Heritage (Bradley et al. 2004). One of the earliest HLCs in a rural landscape was the Peak District National Park project with a specific emphasis on time-depth (Barnatt 2003; Fairclough et al. 1999, 64-65), and for an urban landscape/townscape, the detailed, 'time-slice' approach of the Sheffield project (Sheffield City Council)[23].

In contrast to earlier projects, the introduction of GIS presented a considerable advantage, allowing the extension or correlation of HLC data with other data sets (Fairclough 2002, 277; Turner 2018, 41). Within a GIS, spatial data and interrogable databases can be connected for a detailed analysis tailored to specific questions (Herring 2009, 67). However, the method also introduced limitations and bias. For example, the polygonisation of areas seems to give the impression of clear boundaries between the adjacent areas that do not exist in reality (Williamson 2007, 67). Furthermore, HLC maps could be seen as true representations of landscape, and as objective and neutral. HLC maps could narrow the view when maps representing other aspects of landscapes were not used in tandem, for example, soil maps to get a comprehensive understanding of the environment (Turner 2018, 45; Williamson 2006, 59, 2007, 66). For example, Finch (2007, 377) identified the bias towards the recognition of deer parks and, simultaneous neglect of coverts as a feature of the fox-hunting landscape, which showed the research bias with a focus on the first category. HLC principles admitting this explicitly stated that there is a degree of interpretation and judgement during the process of assigning a character type to an area (Clark, Darlington and Fairclough 2004, 6).

Furthermore, HLC was critiqued for being used to legitimise contested projects and support a political agenda (Finch 2007a, 378; Williamson 2007, 69). However, to be value-free and treat all landscapes equally importantly, HLC principles stated that the method interprets the character in present landscapes, while leaving the decision about the future to others (Bradley et al. 2004, 6). The risk of being interpreted in a certain way to legitimise or support a specific agenda is inherent in the medium. However, maps as support for development and sustainable change in a Western context have a century-long tradition. Important in this regard is that HLC maps must be used within the range and scale of application they were designed for; otherwise,

[23] These areas were defined as study areas for this research and will be presented in detail throughout this thesis.

conclusions inferred from the maps can give wrong impressions (Turner 2018, 46; Monmonier 1996).

Nevertheless, **Chapter 1.2.6** showed the various research projects on HLC and a technical review carried out by Lancashire County Council for English Heritage gave a wide overview of applications of HLC in local authorities (Clark, Darlington, and Fairclough 2004). Furthermore, Natural England suggested inclusion of the historical perspective as 'Historic Landscape Character Assessment' through HLC in England and Historic Land-use Assessment (HLA) in Scotland (Herring 2009, 62) as a useful dataset in Landscape Character Assessments (Clark, Darlington, and Fairclough 2004, 21-26; Countryside Agency 2002, 9-10; Tudor 2014, 27, 50). Also, initially developed in the UK, HLC has been applied internationally and adapted to the various requirements of the countries' specific landscapes (Fairclough 2002, 279).

National coverage of HLC was achieved in 2017 by combining individual projects into a single National HLC dataset available online (Exegesis and Locus Consulting 2017; Natural England 2020). One advantage of this dataset is to function as a '"gateway" for non-specialists into historic resources [...] and in turn make better use of local evidence and expertise' (Exegesis and Locus Consulting 2017, 15). However, the same report highlights the issue that the expert-led HLC design and methodology could discourage laypeople from using HLC and that the original ambition to include terminology and categories from community projects, such as Common Ground, was not realised (Exegesis and Locus Consulting 2017; Fairclough et al. 1999, 12). It has become clear that HLC is not self-explanatory or readily understood without training laypeople and professional sector members (Clark 2003, 78). Also, the generalised, large-scale approach to landscapes allowed the identification of wider patterns. However, this also means that HLC does not (and was not intended to) work on a local level (Fairclough et al. 1999, 5). It is, therefore, questionable whether HLC is useful for local-level decision-making and planning (Clark 2003, 63, 92; Williamson 2007, 69).

Furthermore, one of HLC's 'guiding principles' was to include the views and perceptions of people 'alongside more expert views' and use 'jargon-free' language (Clark, Darlington, and Fairclough 2004, 6). However, HLC was carried out in a top-down approach based on expert knowledge and influenced by the disciplinary traditions of the experts involved (Dalglish and Leslie 2016, 215), adhering to the Authorised Heritage Discourse (AHD) (Smith 2006). The focus of experts in the design of HLC was thus on the material fabric and visible character aspects of the landscape (Turner 2007, 41), not incorporating the perceptions and opinions of people living in these landscapes (Dalglish and Leslie 2016, 215). This issue made HLC challenging to use for non-experts. The missed opportunity to integrate community values has been highlighted as a barrier to a broad uptake of the method (Clark 2003, 45, 75)[24].

In contrast, LCA specifically addresses an area and project and creates a snapshot assessment in time, which is suitable for site or small-scale projects. LCA also has scope for the participation of communities in the project areas to contribute local knowledge and perception. However, as mentioned above, the main assessment work is desk-based and expert-led. Also, while previous projects can be reused, the valuation, particularly the community aspect, which is fluid and dynamic, cannot be addressed when reusing such data.

In summary, while both LCA and HLC provide opportunities to manage the historic environment, no practical way has been found yet to include public perception in HLC and 'meaningful public participation – remains to be addressed' (Dalglish and Leslie 2016, 216; Turner 2018, 45). Limitations and critique of the methods can provide a basis to build on and develop methods that work on a landscape scale and address the 'cultural turn', including people's needs, visions, and values. Learning from and adopting the method's elements will help to develop participatory or bottom-up approaches to foster communication between local people and local authorities. A methodology to evaluate public values would need to be an 'iterative and ongoing process' according to the fluid and dynamic nature of such values (Dalglish and Leslie

[24] See **Chapter 1.2.6**.

2016, 217), hence the need for a method that is resource-effective, using the tools of the digital age (Pink 2016; Kwan and Ding 2008).

2.7 Characterisation and socially sustainable local planning

The challenge for heritage and landscape management is to find the balance between the different views on the development of the historic environment and decide on a sustainable approach to change and continuity (Fairclough 2007, 84-85). Opinions change with the zeitgeist of the time – what is perceived as 'ugly' might be deemed important for the future. Change can be seen as something that has 'uglified' the landscape (Hoskins 1977, 298)[25] or as important evidence of a recent era (Bradley et al. 2004), such as industrial archaeology (Lynch 1972, 49). Fairclough argues that heritage management needs to emphasise this process of evolution, meaning 'change not destruction' which is neither good nor bad – one of the key principles of HLC (Fairclough 2007, 84-87). Heritage is part of everyday life and ordinary landscapes that are imbued with meaning for local people in a sense that 'all buildings and spaces, whatever their age and however modest, make some form of contribution or have value to society' (Worthing and Bond 2008, 1). As such, there is a need to approach change in a participatory, inclusive, collaborative way. Understanding social structures and 'how everyday life practice and places are constituted and how they change' is crucial for this approach (Pink 2012, 149). Pink (2012, 11) suggests that sustainability, challenging environmental problems and enhancing the quality of life can be achieved locally.

The Local Agenda 21 was a United Nations action plan that, while not legally binding, advised on the realisation of sustainable development and emphasised the recognition of social values and local democracy (Chitty 2016; Cinderby and Forrester 2005, 145; Dalglish and Leslie 2016, 213; English Heritage 1997, 2; Gard'ner 2004, 91, Footnote 41; Grove-White 1996, 13). Key points of the agenda, emphasised the role of local government and community engagement (United Nations 1992, Chapter 40), as

---

[25] See also **Chapter 5.**

well as the empowerment of specifically marginalised or vulnerable groups in society, such as children, women, and the elderly (United Nations 1992, Chapter 7, Article 7.4). In response, the UK government passed the Localism Act 2011 (UK Government 2011), transferring specific rights to local governments and encouraging local action with initiatives such as Neighbourhood Plans (UK Government 2011, Part 6,Chapter 3) and Local Lists with Assets of Community Value (UK Government 2011, Part 5: Community empowerment). In the example of Sheffield, local residents could nominate *Assets of Community Value*[26] and additions to the *Local List*[27].

Furthermore, the Social Value Act (2013, updated 2021)[28] was introduced to ensure local authorities acted in the best interest of local communities and for the social and environmental benefit of the public, ensuring social value in the procurement process. The *Social Value Model,* underpinning this policy, follows the United Nations' goals and supports, for example, the fight against climate change with pro-environmental action, 'effective Stewardship' (Theme 3, p. 15), and community integration for well-being (Theme 5, p. 29)[29].

The focus on local places and social value opened new areas for the application of HLC, e.g., Neighbourhood Plans, Village Design Plans, and Conservation Area Appraisals as part of the local planning process which could influence how landscapes evolve (see Clark, Darlington and Fairclough 2004, 41). The Peak District locally designated 12 Neighbourhood Areas with Neighbourhood Plans at the time of writing (PDNPA 2021). The Peak District National Park Landscape Strategy included Village Plans created by residents emphasising their value in the landscape and Conservation Area Appraisals (PDNPA 2009).

---

[26] https://www.sheffield.gov.uk/your-city-council/assets-community-value

[27] https://www.sheffield.gov.uk/planning-development/local-list-heritage-assets

[28] https://www.gov.uk/government/publications/social-value-act-information-and-resources/social-value-act-information-and-resources

[29] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/940826/Social-Value-Model-Edn-1.1-3-Dec-20.pdf

HLC can be a vital planning tool in various conservation and heritage management practices to support sustainable development (Dalglish and Leslie 2016, 213). It was, from the outset, designed to be integrable into a framework of methods, not a stand-alone application for planning (Clark, Darlington, and Fairclough 2004, 11), for example, as part of LCA (Fairclough et al. 1999, 7; Tudor 2014, 14) and Landscape Sensitivity Assessment (LSA) (Tudor 2019, 14, 19). Tudor explicitly described the importance of values held in the local area by local communities and reinforced the importance of assessing these values in advance of planning decisions and that 'value(s) of landscapes and their component parts can be ascertained qualitatively with reference to, for example: Their character and sense of place [and] Their community value' (Tudor 2019, 18). The guide advises using input from local plans to collate these data of locally held values, as well as sources such as LCA, 'Local Distinctiveness studies and other community-produced, place-base documents, community-produced guidebooks, and guides for tourists for example', and refers to Public Participatory GIS as a method of community value data gathering (Tudor 2019, 18). This approach advises an in-depth assessment of locally held values as a snapshot to react to a specific development project, similar to the projects described in **Chapter 1.2**.

However, Arnstein's 'Ladder of participation' (1969, 216, Fig. 2) described public engagement as being 'a little like eating spinach: no one is against it in principle because it is good for you but it is an acquired taste'. Integrating qualitative research into the work of local authorities is challenging because of time and budget constraints. Another is the demand for experts to adapt to new sources of information that are unfamiliar and difficult to assess, as shown in **Chapter 1.2.6**.

Sustainable development depends on the cooperation of local authorities with experts and local communities and the values these groups hold for a specific place. Whether decision-making concerns the location of new residential areas, quarries, routes for railway and road schemes, or investment in the leisure and tourism industry, local authorities must consider the interests of residents and visitors to gain civic trust and cooperation. In reaction to this development for a more socially sustainable, inclusive, and transparent landscape and heritage management several approaches and toolkits provide opportunities to engage local people's social values in the decision-making process.

2.8 Conclusion

The societal shift, which started in the 1960s, was conceptualised in charters and conventions concerning the cultural and natural heritage. While social values were added to the traditional canon of heritage values, the realisation of the aim to include people and local knowledge meaningfully in the tools and frameworks of heritage practice and management is underway, particularly in academic research. The reason for a slower uptake in practical planning and decision-making is the lack of training provided in the heritage profession, as well as challenges to capture qualitative data and translate this into databases that make processes reproducible and comparable. In addition, no appropriate tools have been put forward and used by local authorities (Turner 2018, 39). Change is, however, inevitable not only in view of the development of landscapes but also because of the changing roles of experts and the public in the process of decision-making. Dalglish (2018, 3) argues that '[I]it's time to change the dynamics of that external policy narrative so that the legitimacy of people's involvement in defining and characterising the landscapes they inhabit is accepted as a matter of justice [...] to move on beyond the rhetoric of community empowerment in landscape policy and towards making that empowerment a reality in the interest of sustainable rural renewal'.

There is a consensus that sustainable development can only happen in cooperation with the communities affected – achieved through meaningful involvement and participation and influencing policies on a local level. Increasing acknowledgement and inclusion of people's sense of place and identity and local knowledge as an essential factor in the planning process can create resilient and coherent communities. This inclusive and transparent approach has the potential to benefit communities, increasing the quality of life. Resilient communities can tackle problems such as isolation, xenophobia, racism, and poverty. 'Topophilia' might potentially foster the wish to care for the environment of everyday life, for example, shopping locally, 15-minute city concepts[30] and walking instead of using a car. Heritage professionals and academics are currently trying find new approaches and methods to include public opinion into the planning process to foster an appreciation of the environment. The contribution local expertise can make to the management of the environment is currently still undervalued in practice. Leading from isolated projects and highly complex research projects, an inclusive approach to heritage and landscape management can generate practical applications by effectively integrating people's views into the framework of planning policies. Reinforcing the connection between people and place through a bottom-up approach within local planning can generate appreciation of the everyday places and a wish to care for this environment. Such an approach can help tackle the most pressing problems of the time and create more resilient communities through positive place-making and strengthening the bond between people and places.

[30] https://www.15minutecity.com/

Chapter 3:

Tweets in the Peak: Twitter Analysis – the impact of Covid-19 on cultural landscapes

Published as:

Tenzer, M. (2022). Tweets in the Peak: Twitter Analysis - the impact of COVID-19 on cultural landscapes, *Internet Archaeology 59*. [Online]

DOI: 10.11141/IA.59.6

Abstract

The Covid-19 pandemic had an unprecedented impact on society, with restrictions on socialising and movement during the three lockdown periods between March 2020 and March 2021 (Baker et al. 2021; Institute for Government Analysis 2021). Easily accessible locations offering the typical qualities of tourist destinations moved into the focus of day visitors in periods of restriction easing. The Peak District National Park (PDNP), a cultural landscape comprising historical places, natural beauty spots, and “chocolate box” villages, afforded all qualities satisfying the urge to escape to the countryside. The impact was also felt in the heritage sector, with a noticeable change of visitor behaviour and the relationship between park residents and day tourists (Jones and McGinlay 2020; Sofaer et al. 2021). In order to understand societal change, social media research gives a unique insight into the sentiments, actions, and controversies associated with tourism, Covid-19, and nature conservation. Especially, the open and public nature of Twitter data offers itself for the analysis of large data sets based on specific search queries at specific time periods. For this research, tweets from the PDNP for three weekends in 2019 to 2021 with different restriction levels were collected. Using R and Python, automated processes allow the time-efficient analysis of qualitative information. This project has extended the standard procedures of social media analysis, such as keyword search and sentiment analysis by an emoji analysis and location entity recognition, focusing specifically on cultural and natural heritage. Using Twitter data in a time-efficient process and creating visually appealing outputs may foster an appreciation of the park’s resources and positively influence the behaviour of visitors and residents. Going forward, improving the relationship between people and places will provide background for the management of cultural landscapes and help tackle environmental issues, such as peat erosion from a large influx of walkers, address the climate change emergency, and help ease the controversial relationship between a living and working landscape and tourism.

3.1 Introduction

If the Covid-19 pandemic changed one thing in particular, it was how people interacted, socialised, and generally behaved as a society. Terms such as 'lockdown' and 'social distancing' were introduced into the everyday language - the former even being chosen as one of the Word(s) of the Year 2020 (Oxford Languages 2022b). This selection not only signifies the considerable impact of the pandemic on societal and individual behaviour, it also stands in stark contrast to the Word of the Year 2015: 😂 (Oxford Languages 2022a). The Peak District National Park was one area where this change was felt. The rich and diverse 'cultural landscape' of the national park afforded qualities and opportunities for everyone during various levels of restrictions - from offering places for socially distanced exercise or remote places for relaxation, to vibrant villages for socialising and historical places for intellectual development. The term 'cultural landscape' was coined by UNESCO in 1972, defining a category for World Heritage as an opportunity to bridge the divide between natural and cultural heritage (UNESCO 1972; 1992; 1997; 2021c; n.d.). Never has the diversity and quality of local places been more important than during 'stay local' orders issued during the pandemic.

The change in society also impacted the way academic research had to be conducted and the choice of approaches, methods, and tools safe and appropriate for research in various fields, such as sociology, geography, psychology and the humanities. Qualitative research methods usually applied in these subjects include ethnographic research methods, such as face-to-face interviews, focus groups, participant observation, group work on-site - methods that had one thing in common: personal contact between the researcher and the participants (Low 2002; Madgin and Lesh 2021; Taplin, Scheld and Low 2002; University of Stirling n.d.). Starting research based on such ethnographic methods in 2020 challenged researchers in several ways; for example, ethics and practicalities of methods. In order to adhere to governmental regulations and provide an environment for safe and socially distanced data collection, methods had to be redesigned and extended to remote and passive options.

However, this challenge was not just seen as a stopgap to overcome the challenges of the pandemic and then return to the commonly used methods. This unprecedented time offered new pathways to interrogating and collecting qualitative data, and developments in IT and computational capabilities mitigate the disadvantage of the restrictions on social contact to a certain degree. Furthermore, a variety of social media platforms are currently used by billions of people across the globe producing vast amounts of qualitative data on everyday topics and trends. Facebook, Instagram, YouTube, Snapchat, WhatsApp, WeChat (China), TikTok, and Reddit are just some of the most prominent players in the world of social media. These online platforms offer ways to build communities of special interest groups and connect individuals to people and places - especially in times of restrictions on movement and socialising. Social media research also offers a way to explore trends and sentiments in society, which is, for example, used by academic researchers (Bertrand et al. 2013; Pulido et al. 2018) and governmental organisations (Social Media Research Group 2016).

From the view of heritage management, tapping into the treasure trove of social media as one 'socially distanced’ method for research allows exploring people’s behaviour, sentiment, and connection to heritage places in particular. While physical visits to museums, archaeological sites, landmarks, and travel, in general, were not possible during the most severe lockdown restriction periods in 2020, many organisations and institutions searched for new ways to connect to people. Virtual museum visits (Bianchini 2021; Gutowski and Kłos-Adamkiewicz 2020; Samaroudi, Echavarria and Perry 2020) and outreach through social media, such as the Twitter channel of Chatsworth House or the British Museum, with over 2 million followers, became a new form of cultural experiences, replacing the 'real thing’ and physical museum visits. Lockdown rules in 2020 varied between a full lockdown, including a 'stay at home’ order from 26 March, transitioning into a phase of gradual lifting of measures from 10 May (Institute for Government Analysis 2021).This easing of restrictions gave people more freedom of movement and a slow return to the 'new normal’, including, for example, pre-booked access to venues of English Heritage and the National Trust adhering to social distancing rules and tightly controlled visitor numbers. How did lockdown and restrictions on movement influence people’s behaviour and emotional state? The positive and negative impact of such regulations was explored by a team of researchers at the University of Cambridge (Jones and McGinlay 2020). The method applied in that project was based on an online questionnaire. On the one hand, the report concluded that the Peak District National Park residents enjoyed the quieter periods with lower visitor numbers, leaving more space for the residents. On the other hand, businesses suffered from the decreasing visitor numbers and at the same time, an increase in 'irresponsible behaviour from other users of the Peak District National Park during lockdown' (Jones and McGinlay 2020, 12).

This project further explores the issues and opportunities associated with the pandemic and the impact of changing visitor behaviour on Peak District National Park residents as expressed through Twitter. Tweets covering the period of a full lockdown to the easing of restrictions with a 'stay local' order (Spring Bank Holiday weekend in 2020) provide valuable insights into the public's emotions and behaviour during the strictest Covid restrictions in a typical UK tourist destination. These data are compared to the same bank holiday weekends in 2019 before the pandemic and in 2021, introducing a post-pandemic phase. The tools used and developed in this project focus on Twitter as the primary data source. Part one of this article presents a standard set of analyses performed on the data, which is widely used in social media research, including a keyword search, hashtag and keyword analysis, focusing on sentiment analysis. To provide a deeper insight into the place attachment of people during the pandemic associated with the Peak District National Park, this typical set of social media analyses was extended by new methods of automating the qualitative data analysis in unstructured texts and the entity recognition and extraction process of geospatial information to allow the mapping and visual representation of place information in part two. This article will give an overview of the methods used, present the result of the small-scale study, introduce the automated process of extracting and spatially locating entities in unstructured qualitative data, and give an outlook on the opportunities provided by social media research going forward.

3.2 *Setting the Scene: background and data preparation*

3.2.1 The study area

The Peak District National Park (Figure 3-1) is opportune as a study area for the impact of short-term and day visitors, located as it is within reach of large cities such as Manchester, Sheffield, Leeds, Stoke-on-Trent and Nottingham. Established as the first National Park in the UK in 1951 and location of the famous Kinder Mass Trespass[1] on 24 April 1932, claiming the right of walkers to roam the countryside, the park is an ideal destination to enjoy the views, fresh air, space, rolling hills and vast moorlands. The cultural landscape of the Peak District National Park combines a rich history with areas of natural beauty and idyllic 'chocolate box' villages. Covering an area of 1438km^2 (555 miles^2) and a population of approximately 38,000 residents, the National Park offers secluded places for relaxation, but the varying landscape equally attracts walkers, cyclists, climbers, and photographers (Barnatt and Penny 2004; PDNPA n.d.). The highest point is Kinder Scout, at 636 m. Usually, 13 million people visit the National Park in any one year. At the time of writing, the Peak District National Park had 2900 listed buildings, 109 conservation areas and over 450 scheduled monuments[2].

One of the factors influencing people's outdoor and travel behaviour, apart from the restrictions resulting from the Covid pandemic, may be weather conditions. The weather over the respective weekends varied from cooler and wetter conditions in 2019 (MetOffice 2019) to the sunniest and warmest May on record in 2020 (MetOffice 2020), and mixed conditions on the weekend during 2021, with a significant rise in temperatures in contrast to the first part of May (Farrow 2021).

---

[1] https://www.peakdistrict.gov.uk/learning-about/news/70-years-of-the-peak-district-national-park/the-mass-trespass

[2] See [1].

Figure 3-1: Study area: Peak District National Park (Map created by M. Tenzer, basemap © OpenStreetMap contributors).

3.2.2 Data source and preparation

The wide range of social media platforms offers a treasure trove of information and insights into people's behaviour and sentiment. Accessibility and open-source algorithms for analysis are two reasons why academic researchers and government departments (Social Media Research Group 2016) are using this resource to identify trends, gauging people's sentiments and public opinion for public benefit and policy making. Social media platforms vary significantly in their user numbers and structure. As of February 2022, there are 57.6 million social media users in the UK (84.3% of the population)[3]. Statistics from 2021[4] show, for example, that Facebook had 2895 million subscribers, Instagram 1393 million and Twitter 436 million monthly users and 192 million daily users (Twitter IR 2021, 2). Twitter ranks sixth in the statistic of most used social media platforms in the UK as of 2020[5]. The users are represented by 43% of households with an income up to £14,000, slightly rising to 60% in households with an income over £48,000 with a slightly rising representation from lowest to highest income in the UK as statistics from 2018 show[6]. Twitter, with Facebook, has been found to be the most frequented platform for information on Covid-19 with 12%, based on a survey conducted in 2020 by the Reuters Institute[7] in cooperation with the University of Oxford. As of 2020 data[8], more than 58.5% of Twitter users were male

---

[3] https://wearesocial.com/uk/blog/2022/02/digital-2022-the-evolution-of-the-digital-landscape-in-the-uk/

[4] https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/

[5] See[3]

[6] https://www.statista.com/statistics/611226/twitter-users-in-the-united-kingdom-uk-by-household-income/

[7] https://reutersinstitute.politics.ox.ac.uk/social-media-very-widely-used-use-news-and-information-about-covid-19-declining

[8] https://wearesocial.com/uk/blog/2020/02/digital-2020-the-uk-what-you-need-to-know/

compared to 41.5% female users. A survey from 2018[9] among 1000 interviewees showed that over the last three months of the survey 33% of Twitter users were 15-24 years old, while 51% were 35 years and older. Urban areas were represented by 50% of Twitter users as compared to 37% living in rural areas in the UK, as found by a survey in 2018[10]. The access restrictions, terms of reuse and privacy regulations of social media platforms vary significantly. For example, private and friend groups dominate the closed structure on Facebook, which makes it more difficult for independent researchers to access the information. Furthermore, access to Facebook and Instagram data, included in the newly formed Meta umbrella organisation (Meta 2021), seriously suffered from the Cambridge Analytica scandal in 2016. Nowadays, Meta offers only limited opportunities for independent researchers based on pre-prepared datasets provided by the Meta AD Library and similar pre-prepared statistical data sources (Meta 2022), which led to several issues for researchers (Benesch 2021; Edelson and McCoy 2021; Gibney 2019; Hegelich 2020; Linebaugh and Knutson; Vincent 2021). In contrast, Twitter introduced the 'Academic Research Product Track' in 2021, allowing researchers extended access to Twitter data (Tornes and Trujillo 2022). This research track offers wider research opportunities through the new v2 API endpoints with a historical search option reaching back to the beginning of Twitter in 2006 (Tornes 2021). The academictwitteR package, released in April 2021 (Barrie 2022; Barrie and Ho 2021) was used in this research to collect and store tweets.

As study periods, the Spring Bank Holiday weekends from Friday to the Bank Holiday Monday of the years 2019 to 2021 were defined as detailed below:

(1) First study period: Spring bank holiday/late May bank holiday 2019 (25-27 May). Covid-19 was unknown at this point in time.
(2) Second study period: Spring bank holiday/late May bank holiday 2020 (23-25 May). This was a period of Covid-19 restriction easing in the first national

[9] https://www.statista.com/statistics/278320/age-distribution-of-twitter-users-in-great-britain/

[10] https://www.statista.com/statistics/611892/penetration-of-social-networks-in-the-united-kingdom-by-geographic-area/

lockdown. Restrictions at that time made clear that exercise outside was allowed but a 'stay local' order was still implemented (Baker et al. 2021).

(3) Third study period: Spring bank holiday/late May bank holiday 2021 (28-30 May). The government had implemented the Road out of the lockdown scheme with a gradual easing of Covid-19 restrictions following a third lockdown phase in the UK: 'This means that most legal restrictions on meeting others outdoors will be lifted - although gatherings of over 30 people will remain illegal.' (UK Government 2021).

The three weekends were identified to compare the behaviour, preferred destinations, and sentiments of Twitter users as presented through their tweets. The picture painted on these platforms is limited by individual discretion and personal decisions about content. Nevertheless, the data provided allow valuable insight into people's perceptions, emotional state, and current trends.

3.2.3 General familiarisation with tweet search output

In a first step, familiarisation with the tweet content is crucial for understanding the outputs of the analysis and signposting inefficient and imprecise search queries and issues in the data. It is, therefore, inevitable to manually assess a large chunk of tweets in an initial assessment process. During this process, it became obvious that, for example, the weather station in Whaley Bridge produced a large amount of weather-related posts that produced an additional data load with non-essential information for this research. However, one pitfall of setting too many and too narrow search terms is the potential creation of blind spots that omit valuable information owing to one aspect being excluded from the search query. The decision to make the search as broad and open as possible favoured a reduction of search terms, keeping exclusions to a minimum.

An initial observation of the unstructured texts limited to 280 characters in single tweets showed several components commonly used in tweets: hashtags, emojis, URLs, @ Twitter handles, and other special characters embedded in the textual information. Important for the analysis of unstructured, free tweet content was, therefore, the preparation (extracting several components) and cleaning of the original text (eliminating special characters, white spaces, numbers, and reverting the text to lower case). This preparation enabled interrogation of the data, using Natural Language Processing (NLP), the application of analysis algorithms for hashtag and word frequencies, emoji and word sentiments, and a focus on location information of users and places mentioned in tweets. The following will elaborate on the various analyses performed on the dataset.

The Twitter search yielded datasets of 554, 759, and 698 tweets in the respective periods. The search query contained only three hashtag search key terms (for full search query, see Appendix) to maximise the data collected for the time period: '#peakdistrict', '#PeakDistrictNationalPark', '#peakdistrictwalks'. Retweets, adverts, and keywords prominent in the tweet stream, for example, '@WhaleyChronicle', '@weatherwhaley', '#WeAreWorkingForYou' were excluded from the search.

3.3 Part One: Keyword and Sentiment Analysis

3.3.1 Methodology

3.3.1.1 <u>Hashtag analysis</u>

Hashtags are commonly used across social media platforms to create groups of specific interests summarised by a key term in the form of the # sign and an expression in a continuous character string. Hashtags give an insight into interests and trending topics. The hashtag was invented in 2007 and has ever since been an essential part of social media (Messina 2022; Panko n.d.).

One step of the standard data processing in social media research is the analysis of hashtags. For this research, hashtags were extracted from the main dataset to understand the frequency of hashtags used and their conjunction with the search query's key terms (hashtags) and other hashtags. One focus in this part of the analysis lay in the shift in trends apparent from newly created hashtags at certain times, which would not appear in other years. Another focus was on the frequency of hashtags associated with specific locations.

3.3.1.2 Keyword analysis

For the keyword analysis, word frequencies and word associations were analysed to define areas of interest grounded in the data provided by the tweet content. Word frequencies were explored as a means of defining the trends and highlighting issues apparent from the daily communication of Twitter users at specific times. Word correlation allowed an insight into the connection between various terms emerging from the word frequency analysis. Words featuring high in the word frequency list, for example, 'views', 'place', 'honeypot', and 'landscape', were further explored to identify the specific trends associated at the time of the investigation. Furthermore, terms of specific relevance for this study, for example, 'monument', 'history', and 'heritage', were used for a statistical association analysis. This routine (findAssocs - find associations) is part of the Text Mining Package 'tm' in R (RDocumentation n.d.). The function findAssocs provides an algorithm that consists of two steps. In a first step collocations are assessed and only tweets containing the given word are returned and further analysed. In the next step a correlation threshold will be defined. The correlation score is calculated from the relative number of collocations, where a score of 0 means that the correlated word never appears in the same tweet as the search term, and a score of 1 means that the correlated word appears in all the tweets that contain the search term. Owing to the high variation of words in tweets, a correlation score of 0.2 is considered significant and words matching or exceeding this threshold are assumed to be statistically associated, providing a quantitative result for further qualitative analysis.

3.3.1.3 Sentiment and emoji analysis

Emojis are essential in non-face-to-face communication, of which the social media context is a significant one (Gajadhar and Green 2005). Emojis are used every second on social media, as the real-time emoji tracker[11] on tweets proves. Increasingly, emojis have been used in academic research (Madgin 2021; Novak et al. 2015; Toepoel,

---

[11] https://emojitracker.com/

Vermeeren and Metin 2019). Emojis developed in different parts of the world as a phenomenon based on the same premise, namely to express emotions in electronic messaging and emails, providing the subtext for content and the writer's intention. However, over the years, thousands of emojis were added as icons also expressing objects and actions. The iconic yellow smiley face is said to have been developed by the graphic artist Harvey Ross Ball to raise the morale in the State Mutual Life Assurance Company in 1963 (Stamp 2013). Emojis (literally translated from Japanese as: 'e' for 'picture' and 'moji' for 'letter'[12]) were invented in the 1990s by a Japanese communication firm, and with the development of the standard encoding 'unicode' and inclusion in mobile phone functionality, the cartoon-like images representing faces, professions, trees and flowers and much more have found entry into standard communication (Hern 2015; Unicode Consortium 2020). As a carrier of emotions[13], emojis form a crucial part in conveying sentiment in tweets.

Sentiment is a crucial part of social media analysis. A sentiment analysis allows an assessment of people's attitudes and feelings towards a specific topic and gives an insight into and opportunities to react to changing moods in society. Several studies have engaged in sentiment analysis in social media content and proved the usefulness of the information to improve user experiences, satisfaction, and assess performance levels (Drus and Khalid 2019; Iglesias and Moreno 2020; Neri et al. 2012; Samuels and Mcgonical 2020).

The approaches to sentiment analysis undertaken in this project used the standard sentiment code *VADER* sentiment (Hutto and Gilbert 2014; Hutto 2022; Malde 2020). *VADER* is specifically designed to analyse unstructured text such as Twitter tweets based on word associations, including slang words and emojis (Hutto and Gilbert 2014). Counteracting the limitation on tweet length (initially 140, now 280 characters), emojis compress meaning and content into a single character and increase the content, similar to the use of '#' and '@' symbols as active parts of sentences and links to topics or other users. A text comparison of several packages to analyse

[12] https://www.collinsdictionary.com/dictionary/english/emoji

[13] See[12].

sentiment from tweets has been undertaken elsewhere (Barai 2021), and it has been shown that *VADER* has advantages regarding emoji inclusion and word association. VADER is a rule-based algorithm using a 'gold-standard' lexicon as a basic dataset to establish the sentiment of text. The analysis tool was developed in the US and is, therefore, ambiguous when meanings diverge between American English (AE) and British English (BE). This will be discussed further in Section 5.

Emojis in tweets are encoded as UTF-8. The constant evolution of emojis requires specific steps performed on the data in preparation. This process includes the exclusion of specific UTF-8 modifier codes, for example, hair colour, skin colour, or gender. This step is necessary to represent these correctly in emoji clouds but also to achieve clear and unique counts of emojis for frequency tables, which summarises, for example, smiling faces of different skin colour.

In order to assess the accuracy of the sentiment algorithms used by *VADER*, the tweets were also manually assessed for the prevalent sentiment of the individual tweets. The *VADER* compound score, consisting of the average of neutral, positive and negative scored text components, was rated by me. The aim was to assess the reliability of the automated process and detect weaknesses in the application to the specific dataset. I annotated the original tweets based on the subjective impressions of negative, neutral and positive sentiment and compared the result with the *VADER* sentiments analysis. The annotation of the comparison resulted in True-Positive, True-Negative, False-Positive or False-Negative. The result was visualised in a confusion matrix[14] (Figure 3-7), quantifying the accuracy of the automated process, which is commonly used in machine learning. However, this form of visualisation has been found to be useful for presenting the results of this assessment process.

3.3.2 Results

3.3.2.1 Hashtag analysis results

The hashtag analysis allowed insight into the trends and themes associated with the

[14] https://machinelearningmastery.com/confusion-matrix-machine-learning/

search terms. While the most frequently used hashtag across the years was #Derbyshire, there was a significant change in the year 2020 when new hashtags in reaction to the pandemic were introduced, for example, #lockdown, #covid19, #COVIDIOTS, #doartslockdown, #staysafe, as shown in Table 3-1. Filtering activities, #walking was mentioned in 4 tweets in 2019, compared to 8 in the 2020 list of hashtags, and 19 mentions in 2021, showing an increase in this activity across the years. Activities mentioned frequently across the three study periods were #photography and #hiking. Notably, #staycation, #camping, and #landscape ranked higher in 2021. This may be a direct effect of the lockdown/stay local requirements of the pandemic restrictions, reduced foreign travel options, and a turn to a 'new normal' with a tendency to spend holidays in the UK rather than abroad. In general, the use of hashtags reduced from 740 unique hashtags in 2019 to 560 in 2020 and 686 in 2021 despite an increase in individual tweets.

While it is apparent that locations in towns and villages dominate in 2019, for example, #buxton (not within in the Peak District Park boundary, but included in the #PeakDistrict), #edale, #bakewell featured high in 2019, there is a notable shift from these centres towards more rural locations, such as #mamtor and #kinderscout. This shift is perhaps based on the opportunity for social distancing and because many businesses were still closed or not yet back to their former capacity. In general, #nature seems to appear slightly more dominant in 2020 (5 mentions compared to 3 in 2019) and was more frequently used in tweets in 2021 (11 mentions).

3.3.2.2 Keyword analysis results

The 30 most frequently used words can be seen in Table 3-2. Most striking is the use of the word 'walk' in the year 2020 with 100 mentions, compared to 49 in 2019 and 60 in 2021. Words associated with Covid and restrictions in 2020 included people (89 mentions), symptom (62 mentions), police (48 mentions), Cummings (43 mentions), isolate (36 mentions), drone (25 mentions), lockdown (19 mentions), busy (17 mentions). While positive adjectives rank high in 2019 (good: 5, beautiful: 8, lovely: 13) and 2021 (good: 1, beautiful: 8, lovely: 10), such positive words are not found higher than rank 8 - 'okay' - and rank 19 - 'good' - in 2020. A similar sentiment shift, as seen in the hashtag analysis, seems to emerge, which will be explored further in the next section, focusing on sentiment in particular.

Table 3-1: Top 40 of the most frequently used hashtags across the study

Hashtags 2019	Count	Hashtags 2020	Count	Hashtags 2021	Count
#Derbyshire	28	#derbyshire	22	#Derbyshire	46
#derbyshire	27	#Derbyshire	18	#derbyshire	26
#art	13	#photography	14	#walking	19
#Outdoors	12	#BankHolidayMonday	13	#landscapephotography	12
#Photography	12	#handmade	13	#EnglishTourismWeek21	11
#weather	12	#print	13	#nature	11
#whaleybridge	12	#Art	12	#BankHolidayWeekend	10
#handmade	11	#Outdoors	12	#countryside	10
#peakdistrictnationalpark	11	#Photography	12	#NaturePhotography	9
#print	11	#weather	12	#staycation	9
#wall	11	#whaleybridge	12	#Artist	8
#edale	10	#gift	11	#BankHolidayMonday	8
#gift	10	#wall	11	#TravelPhotography	8
#MyNewTag	10	#BankHoliday	8	#Hiking	7
#Art	9	#Local	8	#landscape	7
#outdoors	9	#peakdistrictnationalpark	8	#NatureLovers	7
#peakdistrictwalks	9	#walking	8	#peakdistrictnationalpark	7
#adventuretime	8	#mamtor	7	#photography	7
#bakewell	8	#EnglishTourismWeek20	6	#art	6
#getoutside	7	#langsett	6	#camping	6
#uk	7	#lockdown	6	#EtsyShop	6
#backtor	6	#landscape	5	#LandscapePhotography	6
#bankholiday	6	#nature	5	#peakdistrictwalks	6
#bankholidayweekend	6	#nikonphotography	5	#walk	6
#buxton	6	#staysafe	5	#WallArtForSale	6
#derbyshiredales	6	#AT	4	#weather	6
#england	6	#BankHolidayWeekend	4	#whaleybridge	6
#highpeak	6	#kinderscout	4	#Buxton	5
#losehill	6	#PeakDistrictProud	4	#cottage	5
#photography	6	#sunset	4	#getaway	5
#ridgeline	6	#AshfordintheWater	3	#highpeak	5
#saturday	6	#buxton	3	#Landscapes	5
#travel	6	#covid19	3	#sunshine	5
#vanlife	6	#COVID19	3	#ad	4
#vwt4	6	#COVIDIOTS	3	#BankHoliday	4
#BankHolidayMonday	5	#DailyChallenge	3	#chatsworthofficial	4
#hiking	5	#DerbyshirePolice	3	#dogfriendly	4
#landscape	5	#doartslockdown	3	#escapetheeveryday	4
#landscapephotography	5	#dog	3	#hiking	4
#mtb	5	#dogs	3	#hols	4

Word associations allow an insight into the relationship of nouns (for example, place names and monument types) with other nouns or verbs (for example, activities) or adjectives (for example, sentiment expressions). For this project, frequently used words from the keyword analysis, such as 'walk', 'see', 'view', 'honeypot', 'landscape', and 'place', were analysed. Focusing on place attachment, place qualities and local heritage, to give useful background information for heritage management, additional terms not included in the word frequency list were also provided to the function, such as 'heritage', 'history', and 'monument'. The word associations can be visualised as a network graph (Figure 3-2).

Table 3-2: Top 30 of the most frequently used words across the study period.

Words 2019	Count	Words 2020	Count	Words 2021	Count
today	62	walk	100	good	73
day	60	people	89	day	67
walk	49	park	86	today	61
good	44	want	73	walk	60
get	36	today	69	get	56
see	35	day	69	just	56
beautiful	33	okay	68	beautiful	55
take	31	symptom	62	one	47
one	31	just	60	lovely	40
look	30	get	58	holiday	37
weekend	30	drive	56	great	36
lovely	29	travel	54	take	35
holiday	28	police	48	live	35
morning	28	weekend	45	hill	33
derbyshire	27	back	44	see	33
come	27	cummings	43	morning	33
edge	26	see	41	home	31
love	26	good	41	place	31
time	26	beach	40	will	30
week	25	can	38	love	29
hill	25	covid	38	view	29
view	24	week	38	park	29
can	23	view	37	time	28
just	22	take	37	trail	28
art	22	holiday	36	visit	27
great	21	ever	36	back	27
place	21	lake	36	look	27
may	21	isolate	36	now	27
bank	20	like	34	edge	27
weather	20	stay	34	can	26

In 2020 the word '**walk**' was solely associated with the word 'dog'. The word '**see**' was associated with 'official', ’stunning' and 'honeypot'. '**View**' found association with 'crystal', 'bowl', 'enjoyable', 'Mickleden', 'stunning', 'Jacobs', 'ladder', 'beyond'. The word '**honeypot**' found association with 'wildfire', 'fire', 'avoid' and 'busy', 'see' and 'forbidden'. '**Landscape**' was associated with 'damage', 'extensive', 'tree', 'nature', 'sheep', 'adventure', 'Peveril' and 'Hathersage'. The word '**place**' found association with 'path', 'canal', 'childhood', 'sanctuary', 'urban', 'restriction', 'Carsington', 'unwind' and 'grateful'. No associations with the word '**history**', '**heritage**', or '**monument**' were found in 2020.

The associations in 2019 were as follows: **history**: corner, fascinate, Tetris ('Tetris' being the name of one of the bouldering areas at The Roaches), glory, build; **heritage**: Bolsover, English, medieval, beautiful, site, castle, countryside, fantastic; **monument**: none; **walk**: iconic, lazy, rugged, route, unplanned, dog; **see**: carry, driver, cyclist, entirely, pleasure; **view**: surprise, stunning; **honeypot**: none; **landscape**: photography, large, wow, vast, bracken, spectacular, church, Padley; **place**: favourite, abroad, adore, Romania, Liverpool, lot, familiar, narrative.

Associations in the year 2021 comprised: **history**: lover, endure, farm, own, plague, guide, stay, tell; **heritage**: English, attraction, destination, national, trust, town, countryside, valley, moor; **monument**: arbor, gib, barrow, henge neolithic, circle, low, stone; **walk**: Longshaw, Grindleford, station, breakfast**; see**: boarder, Rutland, visibility, charismatic, eagle, lynx, marten, wildlife, pine, flourish; **view**: vale, sorry; **honeypot**: none; **landscape**: art, nature, tree, photography, birch, woodland, photo, green; **place**: geographically, Liverpool, vibrant, city, self.

The result of the word association analysis can be understood as a statistical association of the given words and words meeting or exceeding the correlation threshold. This information can be used to identify issues through a follow-up contextual association. A comparison across the word correlation and keywords used in tweets during the study periods shows a notable shift from destinations and travel abroad in 2019 to a focus on local places associated with words of familiarity and safety ('sanctuary’) in 2020. This trend of localism and travel destinations in the vicinity seems to be continued in 2021 with a tendency to explore reopened sites of English

Heritage and the National Trust as well as a significant increase of associated monuments, for example, Arbor Low (henge monument) and Gib Hill barrow, stone circles, and Longshaw (National Trust estate). Notable is, however, also the association of 'landscape’ with 'damage’ and 'honeypot’ with 'wildfire’, which could hint at problems with negative landscape impact factors in 2020.

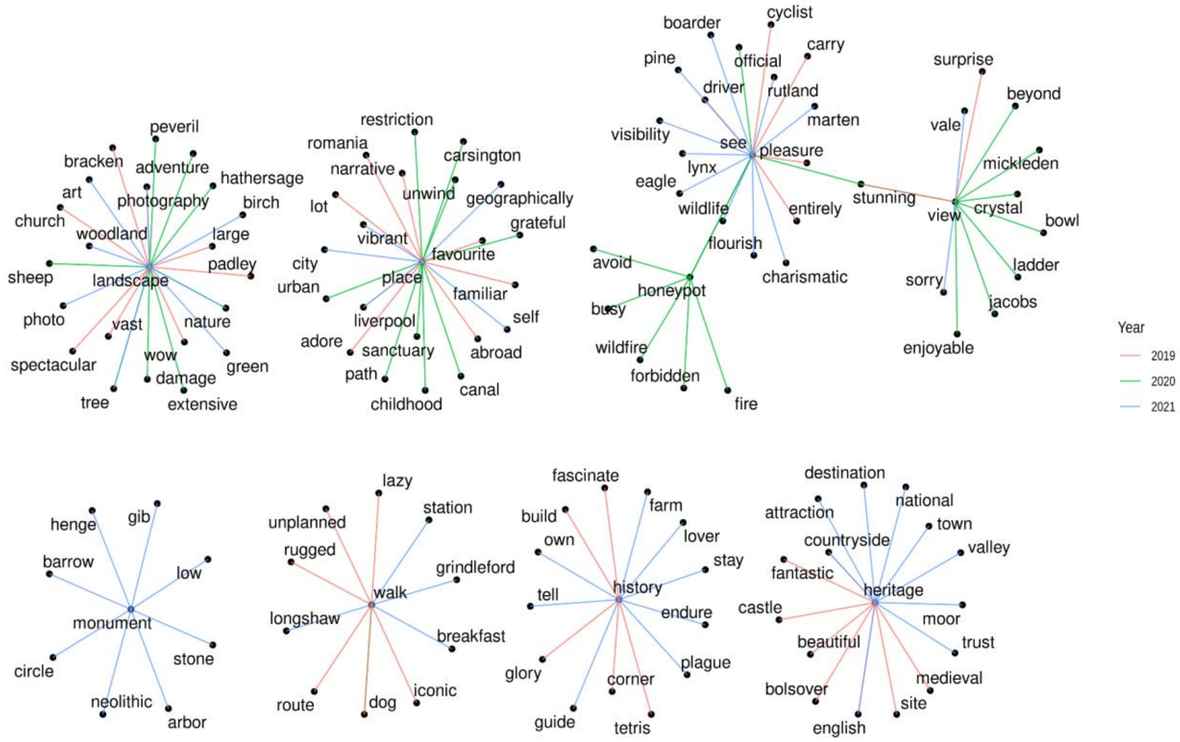

Figure 3-2: Network graph visualisation of word associations based on frequently used words, such as 'walk’, 'see’, 'view’, 'honeypot’, 'landscape’, 'place’, as well as words of special interest regarding heritage management, such as 'heritage’, 'history’, and 'monument’. The colours represent the three years. The visualisation represents statistical associations of the words.

3.3.2.2.1 Visualisation - Word clouds

Word clouds were developed first by the American socio-psychologist Stanley Milgram, visualising landmarks in Paris as text on a map (Cao and Cui 2016, 60; Milgram and Jodelet 1992, 96, Fig. 8.6), but the popularity of this form of visually conveying textual content only came with Flickr in 2006 and the popular tag visualisation for exploring their website (Temple 2019). Word clouds have disadvantages - for example, they do not show word associations or contexts of single words (Cao and Cui 2016, 19), and apply a filter that is either too restrictive (emphasising long words) or too broad (including short words with irrelevant contribution) (Temple 2019). However, advantages of word clouds are the instant understanding of the visualisation without the need for explanation, legends or descriptions. The visualisations are attractive and colourful and allow the representation of keywords of unstructured texts in a compact form (Temple 2019). The word clouds in Figure 3-3 present the most frequently used words in tweets across the study periods in a visually appealing format. The words are not spatially located at the corresponding places of tweet locations.

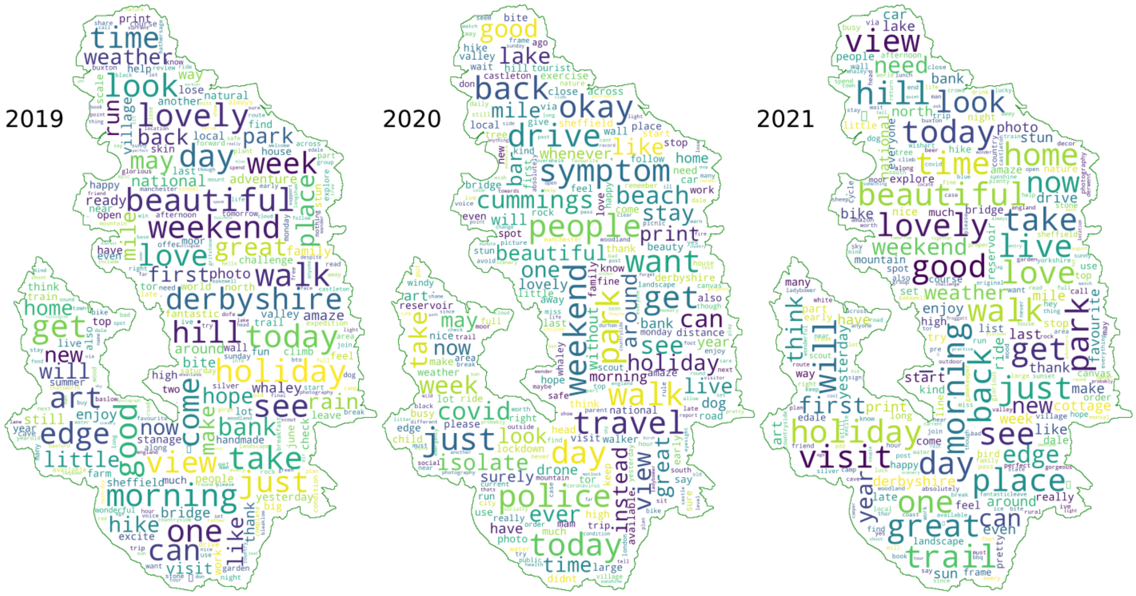

Figure 3-3: Word cloud of most frequently used words across the study period. Disclaimer: The words and their position are not geospatially located within the Peak District National Park, but rather randomly located within the boundary to visualise the summary and frequency of the most used words in tweets related to the area within these boundaries.

3.3.2.3 <u>Emoji analysis results</u>

Sentiment and emoji analysis are closely related since much of the sentiment in tweets is carried by the use of emojis rather than text owing to the characteristic shortness of tweets. We should understand emojis as not one but two categories of pictograms (the western association of emojis = emotion is misleading). Category one emojis expressing emotions, which developed from the earlier emoticons, are ideal for conveying sentiment on its own or in conjunction with text, and *VADER* integrates this aspect for a comprehensive result. We need to distinguish this category of emojis from pictograms used to convey meaning. Emojis, such as a tree, a fruit or boots, replace objects and actions. With regard to sentiment analysis, this categorisation is important, as emotion-emojis are unlikely to change their meaning (a smiling face will unlikely express a negative feeling in future), while object- or action-emojis can change meaning drastically. Within the *VADER* lexicon both categories are replaced by their respective descriptive text and become part of the text that is then analysed for sentiment. *VADER* includes the transcription of 3570 different emojis into text at the time of writing and integrates this into the analysis. Based on this descriptive text and the independent score of 10 individuals (Hutto and Gilbert 2014, 220), the valence score was determined, and all emojis with a non-zero mean score were integrated into the lexicon[15]. This section will present the data of the emoji analysis and present the results of the *VADER* algorithm applied to text and emojis as far as they were known at the time of development in 2017.

The number of emojis used in tweets across the weekends are as follows and shown in detail in Table 3-3.

2019: 110 different types, 267 in total

2020: 118 different types, 313 in total

2021: 121 different types, 337 in total

The most commonly used emojis across the years on rank 1 in 2019 and 2020 and rank 3 in 2021 was the *smiling face with heart-eyes* associated with a positive sentiment. Notable is the increase of negative sentiment expressed through emojis, such as the *angry face* and 🤬 *the angry face with symbols on mouth*. The message associated with this category of negative emotions in 2020 and a complete lack of these emojis in 2019, and for 🤬 with only 2 appearances in the dataset and with only one use, can be interpreted as associated with rising tensions in 2020. Users were positioned on both sides of the spectrum: visitors who felt anger towards the excluding attitude of residents, and residents feeling overwhelmed and angry towards

[15] https://github.com/cjhutto/vaderSentiment

visitors, littering, parking in non-designated areas, overcrowding, and increasing wildfire risks from barbecues. Associated with the latter is the increased use of the 🔥

Table 3-3.: Top 20 of the most frequently used emojis across the study

Emoji 2019	Count	Emoji 2020	Count	Emoji 2021	Count
😍	29	😍	22	☀️	25
🥾	19	☀️	13	❤️	18
❤️	11	👍	11	😍	17
👍	10	🤔	9	😊	13
☀️	9	💙	9	😀	11
😂	9	🙌	8	😎	10
💚	9	🖕	8	🌞	10
⛳	7	👌	7	👍	9
😁	7	🌞	7	😂	7
😊	6	😂	7	👉	7
🤗	5	😁	7	🌿	7
😀	5	😡	6	🐑	7
🍺	4	🤬	6	🌳	6
🐑	4	❤️	6	👇	6
🌿	4	😎	6	🤣	4
👌	4	🌹	6	🥾	4
💕	4	🔥	6	☺	4
🌞	3	💨	5	🐾	4
🐺	3	😊	5	🍻	4
🌸	3	😃	5	💙	4

emoji in the same year. This emoji, within a grey zone between emotion and object, plays a specific role in the issues of the year 2020. Fire, as the descriptor of this emoji, scores negatively in the VADER lexicon[16] (-1.4). Negatively labelled emojis (*fire, angry face*, and *face with symbols on mouth*) were non-existent in the dataset of 2019, increased to 18 in 2020 and decreased again to 3 in 2021. A new introduction of emojis into tweets is the emoji, referring to safety measures associated with the Covid-19 pandemic. As for emojis of category two, the green-leaved branch 🌿 titled *herb* is usually used to represent nature in the main context of tweets and is often associated with the #naturelover. Equal emphasis on nature and the rural character of the Peak District National Park are expressed in the use of the 🌲 emoji, 🌳 emoji, 🌱 emoji and 🐑 emoji. Interestingly, no sheep emoji is used in 2020 tweets.

---

[16] https://github.com/cjhutto/vaderSentiment/blob/master/vaderSentiment/vader_lexicon.txt

Summarising the results of the emoji analysis, the number of negative emojis varied across the years; however, compared to the number of positively labelled emojis, the number of negative emojis stayed low. While still overwhelmingly positive, the year 2020 saw a notable increase in negative emojis and the introduction of emojis associated with the pandemic. The graph in Figure 3-4 shows the variation of negative and positive emojis used across the years based on emojis using only the *VADER* algorithm.

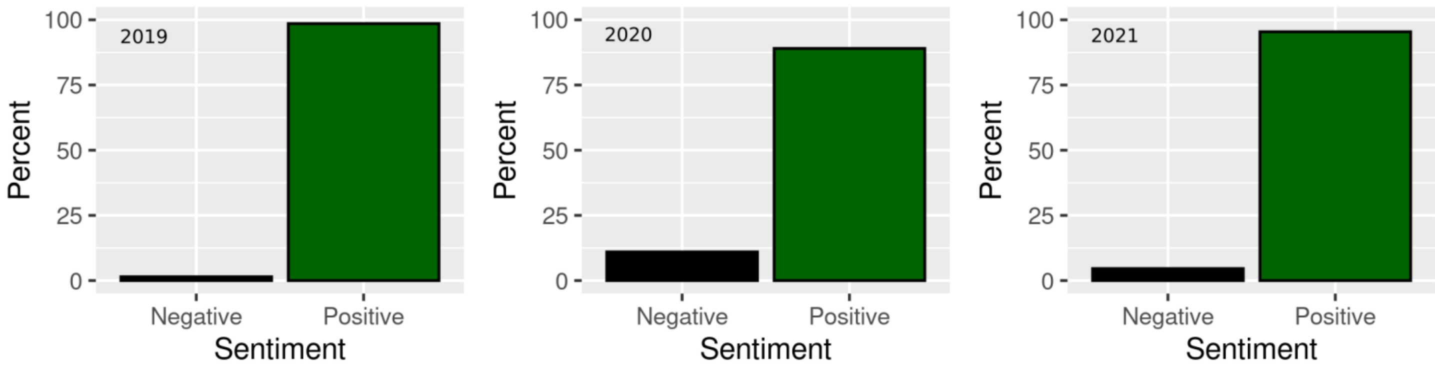

Figure 3-4: Sentiment analysis with VADER sentiment algorithm based solely on emojis. The emoji score is normalised for the sentiment categories of the respective years. Note: the category 'Neutral' has been excluded from the visualisation, as this would skew the tweets with no emojis, neutral emojis and unknown emojis.

3.3.2.3.1 Visualisation - Emoji clouds

Emojis and cloud visualisations have a great potential to convey complex information directly. The format of emoji-cloud visualisation offers a new way to create an instant understanding of the overall situation, trends, and sentiments of large datasets. The emoji clouds presented in Figure 3-5 clearly show the emotional shift in 2020 and can be used to inform and engage the public, similar to the word clouds in Section 3.3.2.2.1, but arguably with a more straightforward and stronger message. The positions of emojis on the map do not correspond with their actual tweet location, and the size of the emojis does not give absolute numerical values. However, this form of visualisation allows instant recognition of the key message and invites engagement with the visualisation beyond passive observation. The key message of a shift towards negative emojis is clearly visible in the year 2020 through the use of distinctly negative icons. Similarly, the smiling face emojis for expressing associations with nature, love and outdoor activities dominate in 2019 and 2021.

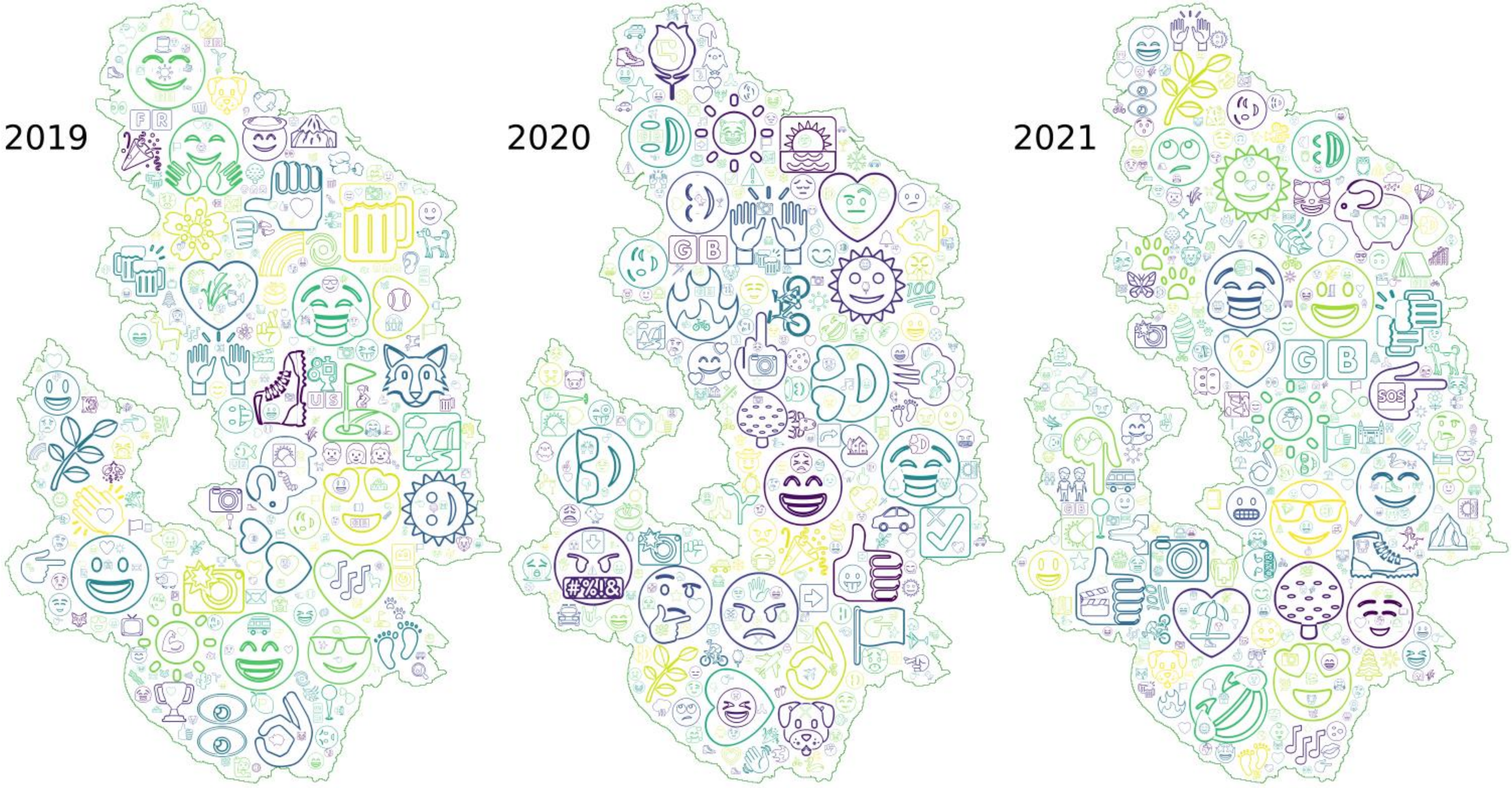

Figure 3-5: Cloud of most frequently used emojis across the study period. Disclaimer: the emojis and their position are not geospatially located within the Peak District National Park, but rather randomly located within the boundary to visualise the summary and frequency of the most used emojis in tweets related to the area within these boundaries.

3.3.2.4 Sentiment analysis

3.3.2.4.1 Results of automated sentiment analysis

This shift to slightly more negative emotions associated with the Peak District National Park shown in the emoji analysis above is mirrored in the sentiment analysis performed on the datasets of all years using VADER sentiment. The score is normalised for the categories 'Negative’, 'Neutral’, and 'Positive’ tweets of the respective years. The analysis result (Figure 3-6) shows an increase in identified negative emotions captured in tweets in 2019 from 6.7% to 14.8% in 2020 and a subsequent decrease in 2021 to 10.3%.

3.3.2.4.2 Reliability of sentiment algorithm

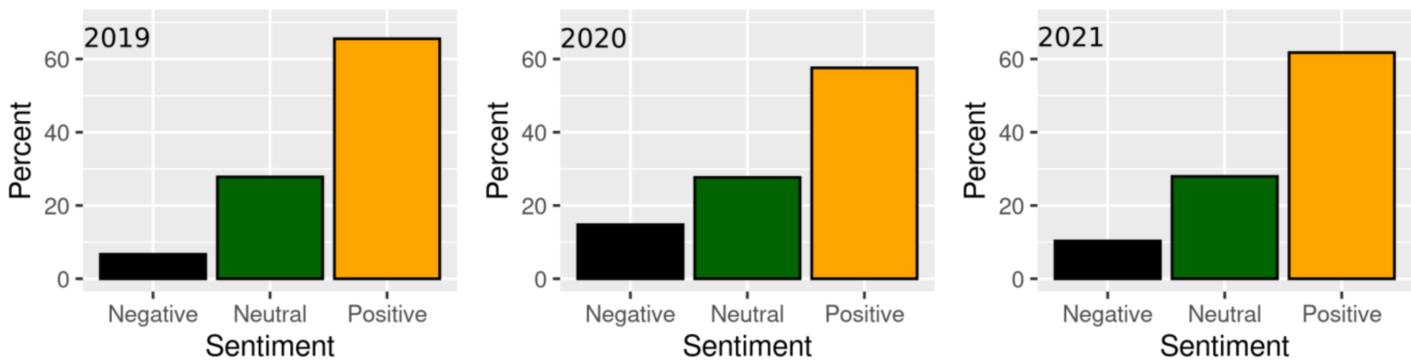

Figure 3-6: Sentiment analysis with VADER sentiment algorithm. The score is normalised for the sentiment categories of the respective years.

To assess the performance of the *VADER* algorithm in extraordinary circumstances as was the case in the year 2020, which has shown a doubling of negative sentiments from the year before, a sample including all tweets from that year has been annotated manually by me. The algorithm seemed to struggle, especially with the sarcasm of Twitter users in response to the Dominic-Cummings-situation[17] travelling from London to Durham during a strict stay at home period. Also, sarcasm of visitors, deterred by

---

[17] https://www.huffingtonpost.co.uk/entry/dominic-cummings-eyesight-barnard-castle_uk_60ae467de4b019ef10e1f1a3

police action and 'stay away' messages from residents, and sarcasm referring to the general Covid situation was misinterpreted by the algorithm. Figure 3-7 shows a confusion matrix for the accuracy as predicted by the *VADER* algorithm and manually assessed by me. The darker blue fields with higher numbers represent higher counts of tweets representing tweets identified as positive, neutral or negative by the *VADER* algorithm on the y-axis and by me on the x-axis, respectively. For example, tweet content identified as positive by both the algorithm and the researcher is positioned in the upper right field of the matrix (dark blue, count: 260), representing True-Positive results. In contrast, tweet content identified as positive sentiment by the algorithm and negative by me is represented by the field in the upper left area of the matrix (mid-blue, count 119), which shows the False-Positive result. The matrix shows a high

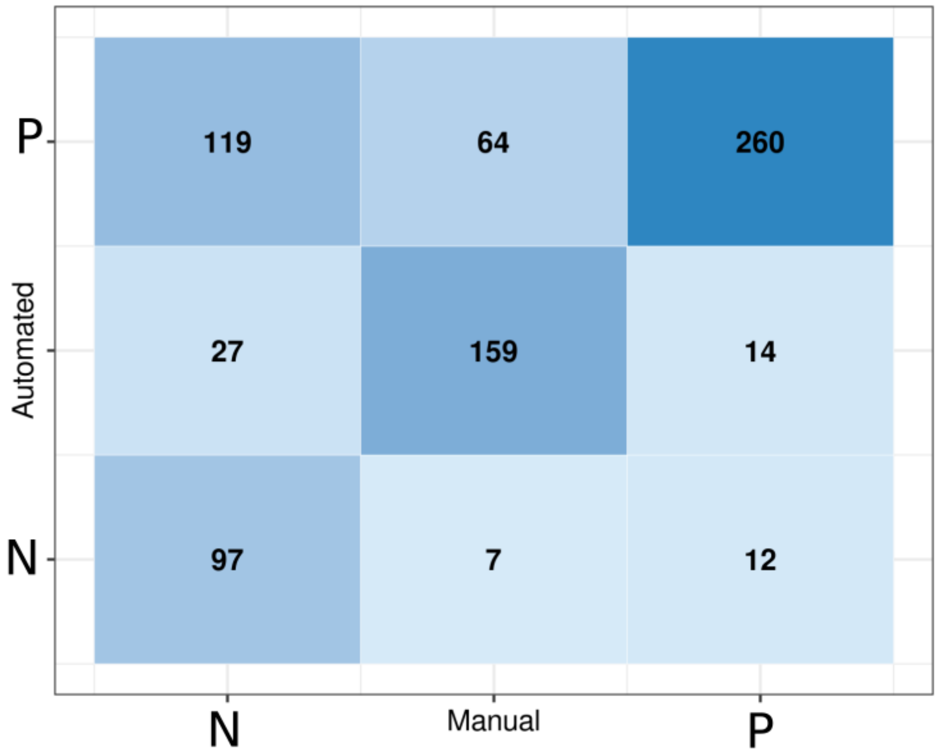

Figure 3-7: Confusion matrix comparing manual and automated sentiment analysis for the year 2020 based on emojis. Bottom left, middle and top right fields show the number of True-Negative, True-Neutral and True-Positive results of the recognition algorithm, respectively. Off-diagonal fields show false matches. Notably the top left field with 119 False-Positives due to misinterpretation of sarcasm in tweets.

agreement of positive results (260) and True-Neutral results (159). However, discrepancies are apparent in the high mismatch of False-Positives (119). The algorithm was evidently not able to identify negative sentiments accurately, in particular if these were expressed using sarcasm and irony.

The accuracy in the case of tweets of the year 2020 can be calculated as:

$$Acc = \frac{\sum_{i=1}^{3} C_{i,i}}{\sum_{i,j=1}^{3} C_{i,j}} = 0.677 = 67.7\%$$

Where C is the count of the respective combination of positive, neutral and negative: C_i for coincidence between automatic and manual assessment, i.e., C_{11}, C_{22}, C_{33} for True-Positive, True-Neutral, and True-Negative, respectively, and C_i with $i \neq j$ for contradictions, i.e., False-Positive, False-Neutral and False-Negative for $j = 1{,}2{,}3$, respectively.

The calculation of the proportion of negative tweets based on the results of manual assessment (Figure 3-7) results, therefore, in 32% tweets with negative sentiment as opposed to the 14.8% identified by the algorithm. In order to assess how this process would adjust the negatively loaded tweets of the years 2019 and 2021, the same process of manual annotation and creation of a confusion matrix was carried out for these years.

The result for 2019 showed that the negative sentiment classified by the

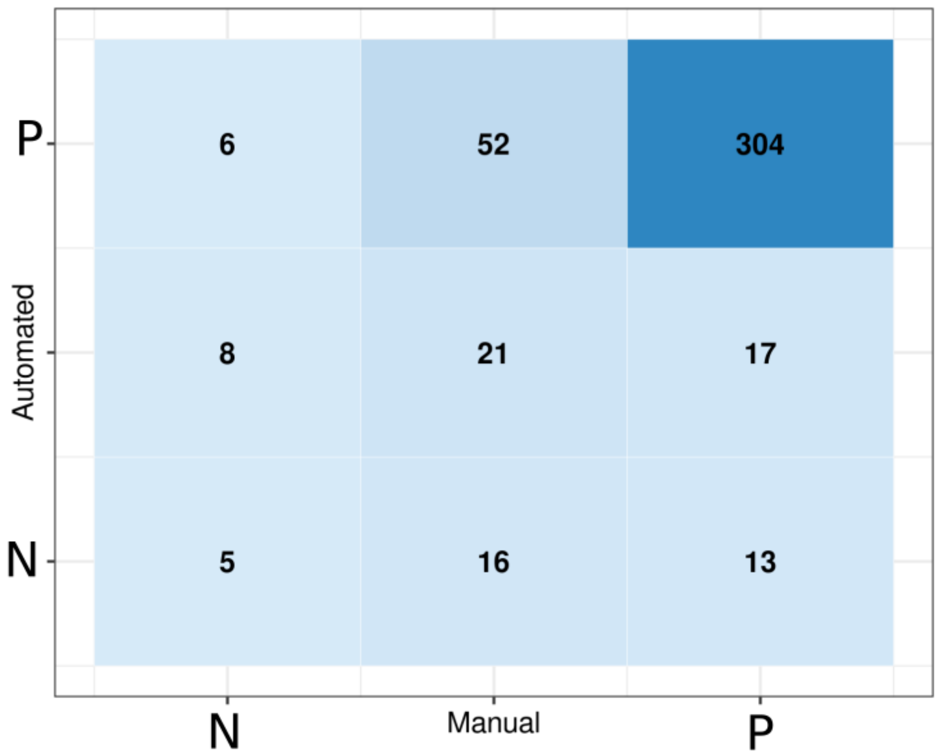

Figure 3-8: Confusion matrix comparing manual and automated sentiment analysis for the year 2019 based on emojis. Bottom left, middle and top right fields show the number of True-Negative, True-Neutral and True-Positive results of the recognition algorithm, respectively. Off-diagonal fields show false matches.

VADER algorithm with 6.7% was higher than when assessed by the manual process with 3.4%. *VADER* also tended to classify more positive tweet content compared to a more neutral manual annotation (Figure 3-8). The manual analysis of the year 2021 data also shows a slightly lower negativity score, with 8.5% compared to the *VADER* sentiment analysis outcome with 10.3% (Figure 3-9). This result shows that the sarcasm of the year 2020 was not detectable by the algorithm, and across the three years, neutral tweets were more often interpreted as positive. Given the (slight) under estimation of negativity in 2019 and 2021, and the under representation of negativity in 2020 by the algorithm, the increase of negativity in 2020 is even more striking in the manual assessment.

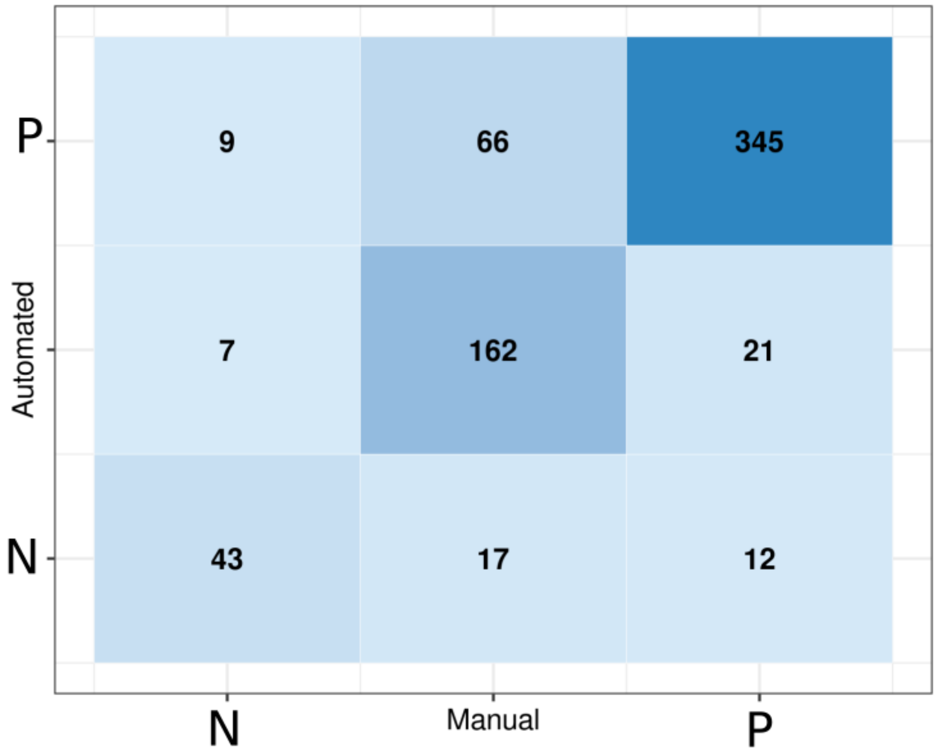

Figure 3-9: Confusion matrix comparing manual and automated sentiment analysis for the year 2021 based on emojis. Bottom left, middle and top right fields show the number of True-Negative, True-Neutral and True-Positive results of the recognition algorithm, respectively. Off-diagonal fields show false matches.

3.4 Part Two: Extracting and Locating Geospatial Information

3.4.1 Methodology

3.4.1.1 User home location analysis - unreliable and fancy

Location data provided by Twitter and by Twitter users consists of three types: 1. geotagged tweet information in the metadata of individual tweets, 2. free-text entries of user home locations and 3. locations mentioned as places in individual tweet texts. Passive geoinformation (1) stored in metadata of tweets strongly depends on the active use of geotagging services, for example, enabling this service and allowing access to location data on the device used for tweeting. On the other hand, location mentioning (3) and information on user home location (2) are open to the user's creativity and the willingness to give more or less precise information about themselves, if any at all. To analyse the latter location information, a frequency analysis can be performed to get a rough idea of where Twitter users are located or where they live.

Geotagged information (1) was found to be unreliable, as the location mentioned in texts and the geotagged information in some test cases did not match. Therefore, no further analysis was performed, including this information. Where user home location data (2) was provided, it was assumed that a summary would give an insight into the adherence of Twitter users to restrictions on movement as issued in the 'stay local' requirement. However, user home location data is imprecise, sparse and inherently unreliable as the input consists of a free text entry. While algorithms determining the home location of Twitter users from the triangulated locations of their geotagged tweets have been developed elsewhere (Mahmud, Nichols and Drews 2014), these are not used in this project since the focus is on place and place attachment rather than on individuals.

Therefore, the main research focus of this project is on the information provided as free text in the tweets (3). This location information was analysed in order to identify places most frequented and talked about during the study periods, which also gave insight into the travel behaviour of visitors to the National Park. Therefore, it is necessary to explore this category further as a component of unstructured text and develop an automated process to overcome the time-consuming manual extraction of geospatial data from qualitative data. This process will be further elaborated on below.

3.4.1.2 General challenges in automated location detection

Geospatial expressions in texts are inherently challenging to analyse, and this applies in particular to the recognition of location entities in unstructured texts such as tweets. The main challenge for identifying and extracting location information from tweets proves to be variations in expressing and naming places. Various levels of 'insider terminology', contractions, and misspelling are a constant issue in the automated recognition of places. However, despite these limitations, an algorithm coded in R and Python allows the extraction of location information from tweets to gain insight into places that matter to people and are trending - for positive or negative reasons - in the wider public.

3.4.1.3 The process of automated location detection in tweets

In order to automatically recognise locations, rivers, points of interest, and historical places, the creation of a gazetteer (a database for geographical information in conjunction with a map in the form of an index or dictionary) was necessary to provide a database for an entity search algorithm. Geospatial information was extracted in QGIS from freely available GIS shapefiles provided by Ordnance Survey (OS) (Ordnance Survey n.d.), the Edinburgh University's Edina Digimap Service (University of Edinburgh n.d.) (note: Edina maps and data are only freely available for research and education purposes), and Historic England's (HE) Listing Data from the National Heritage List of England (NHLE) (Historic England 2022a). The datasets were merged and clipped to the study area. The resulting attribute table was exported to a spreadsheet and cleaned in a further step to delete duplication and exclude some entries, such as pharmacies or surgeries, to focus on places associated with natural or cultural landscapes. Businesses, such as tea shops, campsites or holiday cottages, were retained as they are part of the tourism industry. The cleaned database was re-imported into GIS to extract the location data of all points (Figure 3-10). The resulting gazetteer of places tailored to the Peak District National Park can be used as a standalone dictionary for place entity recognition in order to extract locations mentioned in tweets. The created gazetteer includes over 5000 entries of, for example, rivers, bridges, locations of mines, stately homes, bowl barrows, farms, public houses, churches, towns and villages, rock formations and caves, and walking trails. The comprehensive gazetteer of points of interest and named landscapes visualised in GIS shows a continuous and sufficiently fine-grained coverage of locations across the Peak District National Park. An algorithm

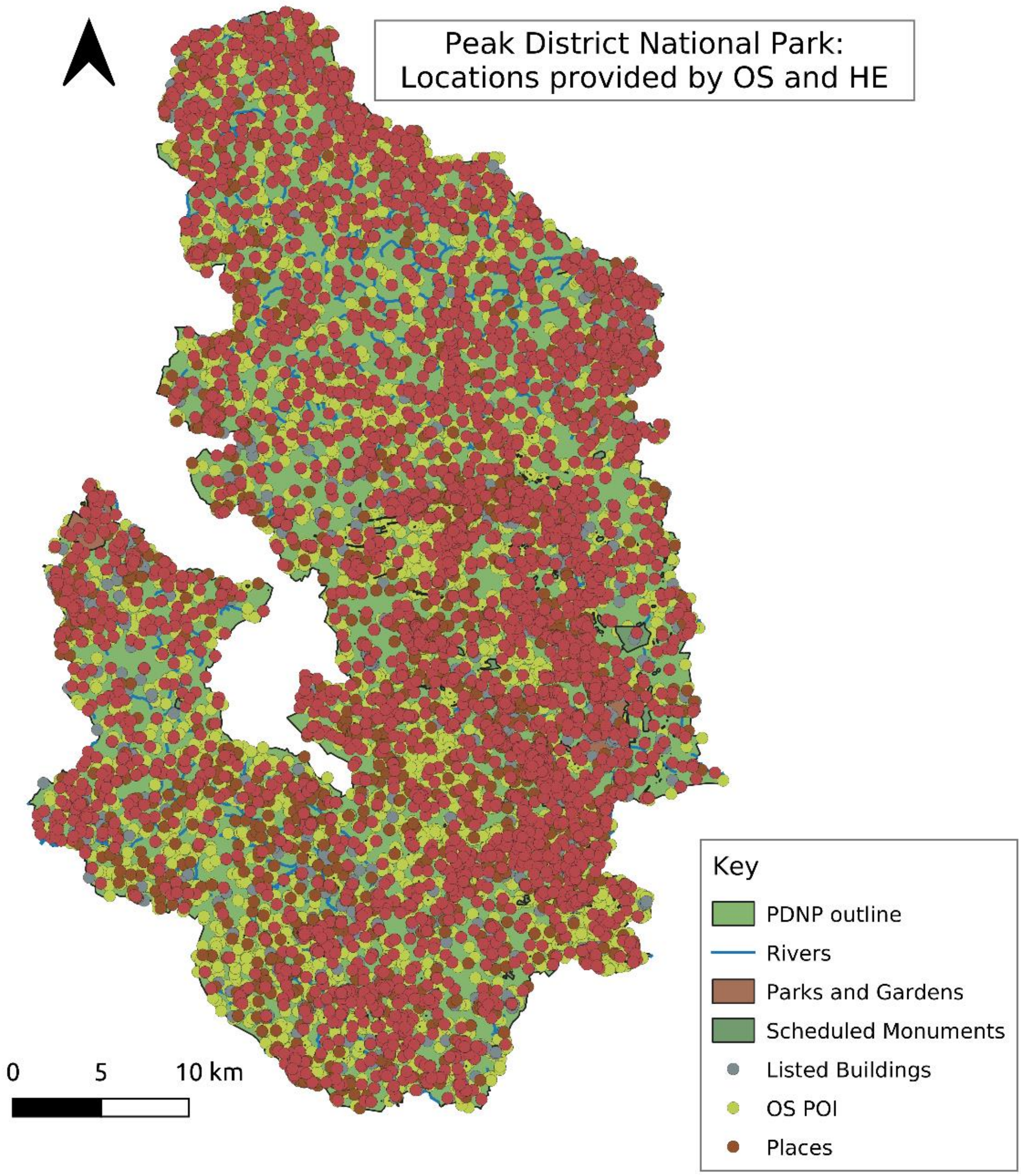

Figure 3-10: Building a corpus of locations in GIS gradually in steps, adding levels of information from various sources, such as rivers, place names, points of interest from Ordnance Survey data, or historic information from Historic England data (Image: M. Tenzer).

implemented in Python compares the gazetteer with the unstructured tweet texts and builds a separate dataset of location entities for further analysis and visualisation. The algorithm extracts single word locations as well as compound terms. As identified during the familiarisation process with the data, the challenge of fuzzy and imprecise location-naming in tweets had to be overcome by allowing the detection of, for example, incomplete location descriptions, such as 'Kinder' for 'Kinder Scout', but also detect the compound term, for example, 'kinderscout' without counting the variants independently when mentioned together in one tweet. This method allows the creation of a unique set of distinct features and places in an area similar to Named Entity Recognition (NER) as part of NLP. However, the method detects locations with higher accuracy through the area-specific corpus of distinct locations.

3.4.2 Results

3.4.2.1 <u>Results of the location analysis</u>

While this research did not go so far as to track down the users' locations through the algorithm developed by Mahmud and his team (Mahmud, Nichols and Drews 2014), the free text self-identification of the users' location gave a rough idea of where the tweeters came from during the lockdown phases (home location of users entered by the users). Did the *stay at home* order work? Did people adhere to the restrictions? Some entries of user location consisted of incomplete, fancy and funny entries and very general spatial information, for example, 'United Kingdom' or 'No boundaries'. Nevertheless, the user location information was analysed on the frequency of occurrence and provides insight into the home location of tweeters (Table 3-4). The analysis shows that across all years, the majority of tweets associated with the Peak District National Park originated from users based in the National Park itself or the surrounding areas of larger cities within reach, for example, Sheffield, Chesterfield and Manchester. During the years 2019 and 2021, user locations further away appear as well, for example, Nottingham, Derby, Wales, Hampshire and Rochester. In 2020, the stay local orders appear to have been followed by the majority of visitors to the National Park, evidenced by the user locations mostly located in or surrounding the park, with some exceptions such as tweets from London and the East Midlands. It is, however, important to note that a tweet from one of these more distant locations does not necessarily mean that the person was actually in the park; tweeters could have just mentioned the park in the text. This was the case for tweets where the tweeter was, for example, longing to come back or missing the visits to the National Park.

Table 3-4:User location of tweets about the Peak District National Park with the 25 most frequently used location.

Users 2019	Count	Users 2020	Count	Users 2021	Count
Sheffield, England	39	Sheffield, England	46	Sheffield, England	34
United Kingdom	17	United Kingdom	45	United Kingdom	26
Whaley Bridge, Peak District	12	Derbyshire	25	UK	16
UK	11	UK	15	East Midlands, England	15
England, United Kingdom	9	Sheffield	13	England, United Kingdom	13
Manchester, England	9	England, United Kingdom	12	Manchester, England	12
Peak District	9	Whaley Bridge, Peak District	12	Sheffield	12
Redcar, North East England.	9	Peak District, Derbyshire, UK	11	Buxton, England	10
Derbyshire	8	Peak District, UK	9	Derby, England	10
Nottingham	7	Manchester	8	High Peak, Derbyshire	10
Rochester, South East	7	Nottingham, England	8	Sheffield, UK	9
Sheffield	7	Derby, England	7	Chesterfield, England	7
London	6	Peak District	7	Derbyshire/Lothian	7
London, England	6	Holmfirth	6	London	7
Peak District, UK	6	London, England	6	London, England	7
Manchester, UK	5	Manchester, England	6	Derbyshire	6
Matlock, England	5	The heart of Yorkshire.	6	Derbyshire - Peak District	6
Bakewell, England	4	England	5	Whaley Bridge, Peak District	6
Hampshire	4	London, UK	5	Peak District	5
Hayfield, England	4	🌎 No boundaries 🏳️‍🌈	4	Bakewell, England	4
Northwich, England	4	Buxton & Sheffield	4	Chesterfield ,Derbyshire	4
Nottingham, England	4	East Midlands, England	4	Matlock, England	4
Peak District and north London	4	England, UK	4	Nottingham	4
Ashbourne, Derbyshire, UK	3	High Peak	4	Wales	4
Bakewell	3	Lincolnshire	4	Yorkshire.	4

In summary, it can be noted that the analysis of user home location is not reliable. Location information from tweet text metadata does not necessarily match the location mentioned in the tweet. In contrast, extracting and geospatially locating places mentioned in tweet texts can offer more reliable data on locations visited by Twitter users, highlighting trends as well as residents' and visitors' behaviour and sentiment associated with specific places in the national park. The automated recognition and analysis of locations in tweet texts was further developed as part of the unstructured text analysis and will be elaborated on in the next section.

3.4.2.2 Results of location entity recognition

The automated process using the gazetteer developed in Section 3.4.1.3 recognised

115 individual places mentioned 332 times in 2019. The weekend in 2020 identified 118 different locations with 237 mentions. This 2020 result was also manually evaluated, the dataset annotated by me, and tested against the automated process described in Section 3.4.1.3. The weekend in 2021 yielded 179 unique places with 406 mentions in individual tweets.

3.4.2.2.1 *Reliability and validity of location recognition*

A subset of data was analysed manually to assess the reliability of the location entity recognition process. The 759 tweets of the bank holiday weekend 2020 were used to compare the performance of the code and the completeness of the location database created in the first step of the process. The manually evaluated list of locations

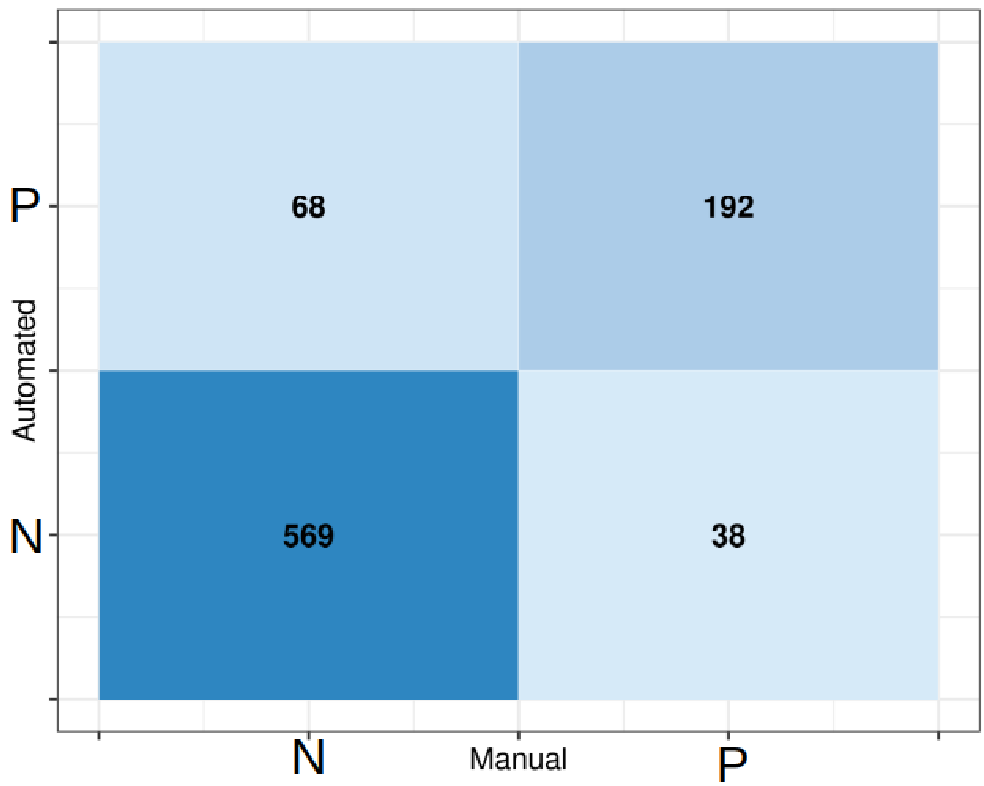

Figure 3-11: Confusion matrix, comparing manual and automated place recognition. Bottom left and top right quadrants show the number of True-Negative and True-Positive results of the recognition algorithm, respectively. Top left and bottom right show False-Positive and False-Negative matches, respectively.

provided the gold standard for the comparison. Results from the manual and automated processes were subsequently evaluated, giving the True-Positive and True-Negative results with corresponding outcomes and False-Positive and False-Negative results where the automated process produced an error. In total, 569 True-Negative (no location mentioned in the tweet) and 192 True-Positive results (location in both processes identical) were recorded. The weakness of the algorithm is shown by the 68 False-Positives (location wrongly identified by the algorithm) and 38 False-Negatives (location not identified by the algorithm). Figure 3-11 shows a confusion matrix visualising the comparison of the algorithm.

The accuracy of the automation algorithm can be calculated using the equation:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} = 0.876 = 87.6\%$$

With nearly 88% accuracy, the algorithm using the area-specific location data proves efficient in recognising places mentioned in tweets. Extracting this information from qualitative, unstructured text is necessary to map the locations as hot spots for activity, which can subsequently be visualised in GIS.

3.4.2.2.2 Visualisation – heatmaps

GIS maps offer increased opportunities for interrogation and engagement with the data. The great advantage of GIS maps is their flexibility, dynamic quality, and ability to connect databases of varying complexity to spatial expressions. Visualising the content of unstructured text, such as tweets, allows insights on various levels and topics, for instance the most frequently mentioned places or activities or problems associated with visitors and tourism. The geospatial data extracted from the tweets are visualised in GIS in the interactive web map shown in Figure 3-12 as point features and hotspots. The web map was created using QGIS v.3.16 and exported to OpenLayers using QGIS2Web plugin[18]. Point data shows the location name and counts of mentions in

---

[18] https://plugins.qgis.org/plugins/qgis2web/

tweets as numbers in a pop-up field. To visually present the variation across the locations based on the weight (numbers of mentions), heatmap layers were created, showing the intensity of frequency. The years 2019 (green point data, green heatmaps), 2020 (red point data, red heatmap), and 2021 (blue point data, blue heatmap) can be used individually and independently as overlay maps to allow interrogation of the various factors. Zooming in on specific locations allows an intensive exploration of the study area, highlighting areas of high impact on landscapes but also access roads, parking issues, and locations requiring individual and specific support on different levels. Owing to the nature of these maps, they can be used intuitively, not just by practitioners and researchers but also by the wider public, to interrogate the landscapes and places of interest in detail.

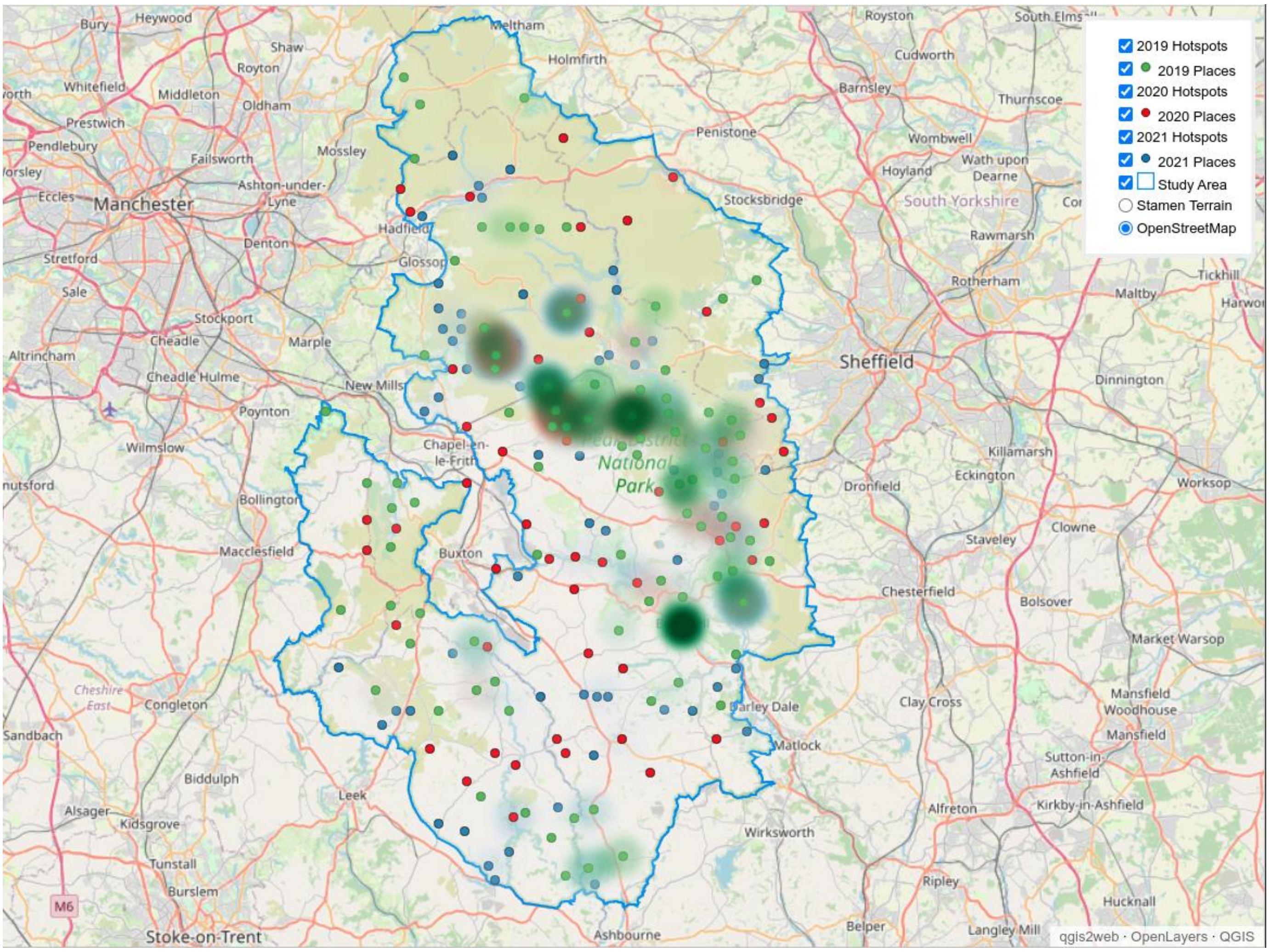

Figure 3-12: QGIS2Web OpenLayers map showing point data and heatmaps of locations mentioned in tweet texts of the Spring bank holiday weekends 2019 to 2021. Base maps used in this project provide the background for orientation and navigation of the map. For map use, please zoom in for details, and switch layers on and off in the top right corner menu (interactive map in online version only).

The interpretation of the maps will be discussed further in Section 3.6. However, the following will detail some of the hotspots featuring prominently on the map. The characteristics and qualities of the locations highlighted on the map vary, including areas of natural, historical and communal features. Bakewell in the White Peak[19], a historic market town on the River Wye, represents a typical 'chocolate box village' character with 17th-century architecture, country fairs and markets, shopping opportunities, as well as the famous Bakewell tart, traditional well-dressing and a carnival. Various features in the town are scheduled monuments or listed buildings[20], such as the Bakewell Bridge, Lumford Mill, the motte and bailey castle and All Saints Church with Anglo-Saxon crosses. Similar but smaller places appear in tweets located in the Dark Peak, for example, Edale, Hope and Castleton. These small villages and towns are situated among rolling hills, offering shopping opportunities for the famous Blue John stone found in the caves around Castleton. Edale marks the start of the Pennine way, a long-distance National Trail. Castleton is situated at Winnats Pass and the 'Broken Road', the A625, a torn and twisted road built on a moving geology typical for the surrounding mountains. Historically, Castleton lies in the shadow of Peveril Castle, the 11th – 14th century tower keep castle. Mam Tor[21], a Bronze Age hillfort of national importance lies just beyond Winnats Pass and offers wide views across the valleys and along the ridgeway, its other end connected to Lose Hill - another favourite viewpoint across the valleys. Hope lies close to the Roman fort 'Navio', one of the few traces of Roman life in the Peak District. These areas are favoured by visitors and residents for the natural beauty, stunning views, and opportunities for various outdoor activities, such as cycling, walking and paragliding. The High Peak comprises Buxton (which is not part of the Peak District National Park) and part of the Dark Peak in the north of the Peak District National Park. Included in this administrative area of the Peak District is the Kinder Plateau, with Kinder Scout being the highest peak in the park, a National Nature Reserve and a Site of Special Scientific Interest (SSSI) formed of

---

[19] http://www.derbyshireuk.net/bakewell.html

[20] https://magic.defra.gov.uk/MagicMap.aspx

[21] https://heritagerecords.nationaltrust.org.uk/HBSMR/MonRecord.aspx?uid=MNA112487

'blanket bog and sub-alpine shrub heath'[22], an ideal location for bird watching and walking (see also Section 6). Historically, the area is associated with the famous Kinder Mass Trespass[23] in 1932 - a protest for the right to roam the countryside freely. Jacob's Ladder leads up to Kinder Scout and is part of the Pennine Way and connected to the scheduled packhorse bridge and packhorse route across the Pennines. North of Kinder Scout lies Bleaklow Hill and Bleaklow Stones, known as the site of a plane crash in 1948, complete with the wreckage of a US Air Force bomber. Also favoured for its views by walkers and the rough gritstone formations by climbers, Stanage Edge features prominently in tweet texts. Relics of the former millstone and grindstone production still dominate this area and features as the logo of the park. Last but not least in importance for the Peak District are the great country houses of the park, of which Chatsworth House[24] is one of the most prominent. The Grade I listed house[25] with its history starting in the 16th century is home of the Duke of Devonshire 'passed through 16 generations of the Cavendish family'[26]. With regular events in the spacious surrounding park, Chatsworth attracts visitors and residents for a day out. Together with Stanage Edge, Chatsworth featured in film productions, such as 'Pride and Prejudice' and there is a strong connection to Jane Austen and the Brontës in the Peak District National Park. The variety of natural, historical and recreational qualities of the Peak District is reflected by the various locations mentioned in tweets and visualised in the hotspot maps.

3.5 Limitations and Researcher Bias

Various factors limit the generalisation of the results. The data source was limited to a

[22] https://www.gov.uk/government/publications/derbyshires-national-nature-reserves/derbyshires-national-nature-reserves#kinder-scout

[23] https://www.peakdistrict.gov.uk/learning-about/news/70-years-of-the-peak-district-national-park/the-mass-trespass

[24] https://www.chatsworth.org/visit-chatsworth/chatsworth-estate/history-of-chatsworth/

[25] https://historicengland.org.uk/listing/the-list/list-entry/1373871

[26] https://www.chatsworth.org/visit-chatsworth/chatsworth-estate/house/

dataset of tweets by Twitter users. Further analysis of other social media platforms was not possible owing to time constraints, limiting investigation of aspects of the structure of other platforms, and ethical approval considerations. A more extensive dataset would have provided more scope to test the application. Because of this limitation, the sample was not representative in a broader sense. However, as surveys associated with heritage are usually limited to visitors to historic sites (English Heritage 2000, 2) or people interested in the subject (National Trust 2017), an advantage of Twitter data is that the dataset reaches beyond the 'usual suspects'. In this respect, the data provided insight into a more variable and diverse group of people.

Limitations of the method are constrained through the accessibility and availability of datasets provided by services such as Ordnance Survey (OS), with finer-grained geo-information only available to commercial or governmental organisations, for example, the '1:50 000 Scale Raster map with Gazetteer'. For educational purposes, these datasets can be acquired through Edina. OS maps provide databases with a searchable data structure allowing data integration. Other datasets, for example, Historic Environment Records (HER) and HE datasets, are difficult to integrate without time-consuming curation of the data owing to inconsistencies in, for example, data entries and unsuitable data structure, which makes querying the data extremely challenging if not impossible. To overcome this problem, datasets were manually adjusted during the compilation of the database for this project. The human component in this research, introduced through manual cleaning of the location database and manual annotation of the test dataset, cannot be neglected. For instance, decisions were made whether to include or exclude locations in the corpus, such as disregarding points that were considered irrelevant to the research aim, for example, pharmacies or surgeries. However, the results were assessed and reviewed several times to improve the algorithm and the database, creating a comprehensive compendium of locations in the study area. In a similar way, the sentiment analysis was assessed by me, introducing a certain level of subjectivity. However, I was fully aware of bias and subjectivity during the analysis phase, and this kind of limitation is part of all research projects that involve a human component. This testing and evaluation process, once finalised, can save time in the subsequent processing of qualitative data in the specific area and, therefore, justifies the effort and introduction of researcher bias/subjectivity.

The visualisation created for this project comprises a map showing point data and heatmap areas, which currently provide limited additional information. However, more details included in the database and additional aspects, for example, sentiment information connected to locations, would provide a broader picture of various aspects of landscape impact and use. Equally, there was no distinction between natural and cultural heritage aspects of the diverse locations. On the one hand, I do not find this a valuable distinction regarding using these maps as a management tool for cultural landscapes in the broader context. On the other hand, this approach aligns with the notion of managing landscapes as a continuous, value-free and complex ecosystem comprising the natural and cultural components. This approach connects natural and cultural features that form cultural landscapes and addresses the various requirements and challenges of mixed landscape form, such as human perception and demands as well as nature and wildlife protection.

The accuracy of the method to automatically identify and extract location data was limited by the ambiguity of location names. For example, 'Hope' as a place name in the Peak District was several times confused with the verb 'hope' that was used regularly in conjunction with Covid-19 and the behaviour of visitors. Another area of improvement of the method lies in the regular use of abbreviated place names in the colloquial speech common in tweets. For example, for 'Kinder', short for 'Kinder Scout', the algorithm recognised both terms as separate entities but counted these as a single entity, but less common abbreviations cannot be identified by the algorithm. Furthermore, misspelt place names contributed to skewed frequency tables and incomplete map overviews. The method of geospatial entity recognition is accurate to almost 88%. Therefore, 12% of the identified locations are wrongly classified - either as 'False Positive' or 'False Negative'. The visual representation of locations mentioned in tweets is, subsequently, only reliable to a certain degree. However, as the main aim of the method is to recognise broader trends and sentiments of groups and not to map individual behaviour and movement precisely, the algorithm proves to be sufficiently accurate.

As noted in Section 3.3.1.3 the results of algorithms applied within the field of rule-based text analysis heavily depends on the underlying database. In the case of VADER sentiment this lexicon was developed in the US, which leads to issues of word sentiment scores based on word meaning. Variation of word meanings between American English (AE) and British English (BE) can lead to ambiguities that should be acknowledged. The word 'quite’ was tested as an example of how severe the difference is overall when used as a modifier, as in, for example, 'quite good’. While numerous meanings in common use are listed in the Oxford Dictionary[27], the meaning in this case can tend towards 'very good’ (AE) or 'somewhat good’ (BE). The test of the algorithm showed a variance of: 'quite good’ and 'very good’: 0.4927, 'somewhat good’: 0.3832, 'good’: 0.4404. This represents a slight variance in the compound result; however, this would not notably change the overall result of the analysis and can, therefore, be discounted.

Furthermore, the text analysis focused on keywords, their frequency and correlation. While it is possible that the meaning of words can change when viewed as bi-grams (two-word term) or tri-grams (three-word terms), this research did not use this method as the focus. This research aimed to provide methods for a practical and easy approach to visualising textual information, and for this purpose the analysis only considered uni-grams.

3.6 Discussion

The standard social media analysis processes in R and Python enable a thorough interrogation of Twitter data collected from three Spring bank holiday weekends in 2019, 2020 (with Covid restrictions in place), and 2021. Steps in the analysis process include keyword and hashtag analysis, which provide the most frequently used terms for a comparison across the study periods. The results have shown the introduction of new words and hashtags associated with the pandemic. A shift in behaviour and activities and variations in locations mentioned frequently across the study periods has

[27] https://www.oed.com/dictionary/quite_adv

also been noted. Overall, the analysis gave insight into trending topics and issues at respective times. Working with the tweet texts allowed the identification of the most frequent words and word associations with these, as well as with heritage-related terms, such as 'history’, 'heritage’ or 'monument’, providing valuable background information for heritage and resource management of the national park. The study has shown a trend towards local heritage sites following on from 2020, which was dominated by Covid restrictions to movement and socialising. However, the results have also shown that the statistical word association can highlight potential issues and tensions that can be further analysed contextually for planning and action in landscape and heritage management.

One of the two main aims of this study was to assess and present a methodology for sentiment analysis with a focus on visually informative and appealing outputs that engage the public as well as provide background for local authorities. Sentiment analysis, as an automated process, continuously struggles with the nature of unstructured text, as presented in the form of Twitter tweets in this research. Human sarcasm, double meanings, and irony are not sufficiently identified by automated processes yet. Additionally, automated processes and machine learning introduce human subjectivity when preparing training datasets and corpora. The manual analysis of the dataset for 2020 made it obvious how subtle and variable meanings are expressed. The sentiment analysis and emoji visualisations highlighted a notable increase of negative sentiment during the pandemic and its relaxation in the following year. Overall, however, the positive aspects and attitudes towards the National Park were dominant in all three study periods. Emojis proved to be well suited for conveying sentiment, as shown by the emoji cloud visualisations. A separate test of emoji-only sentiment showed that this part of unstructured text can influence the overall sentiment score of a tweet. The integration of emojis in the sentiment analysis was crucial owing to the frequent and increasing use of the symbols in social media and the advantage of the instant understanding of meanings through emojis. The second main aim of this article was to present a new method of location data extraction from unstructured text. The geo-tagged tweet location and the free-entry user location in Twitter metadata have proved to be less useful for an analysis focusing on place attachment. In contrast, the method for automated recognition of locations mentioned in the tweet texts, developed for this research, proved to be efficient, highly accurate, and effective. The automated recognition of locations was achieved by creating a gazetteer of locations provided and an algorithm that is able to extract the places, landmarks, buildings and natural features by comparison with a dictionary from tweet texts. While comparable algorithms in text mining software already use a similar process (Named Entity Recognition), the innovative aspect of this research lies in the method of compiling a gazetteer from existing databases providing a framework for fine-grained entity recognition in specific areas. The accuracy of corpus and algorithm of almost 88% provides a sound basis for a sufficiently detailed visualisation in GIS. The visualisation shows a notable shift from favoured locations before the pandemic, mostly focused on 'honeypots' and 'chocolate box villages', such as Bakewell and Castleton in a wider distribution across the study area, to natural beauty spots, such as Mam Tor and Kinder Scout during the pandemic restrictions (Figure 3-13 and Figure 3-14). The reasons are most likely the ability to socially distance in these areas and a welcome change of scenery for city dwellers. The temporary closure of businesses in villages and towns across the Peak District during lockdown shifted the most frequented visitor locations from villages and historic buildings towards the open countryside. Being confined to local areas or even the home over a period of weeks raised the awareness of the qualities of places in the vicinity that offered a change of scenery, an opportunity to breathe, increasing the feeling of escaping to the countryside. These conditions led to a high influx of visitors, with no economic advantage for the region. Associated disadvantages, such as overcrowding, damage to moors from barbecue fires and parked cars in undesignated areas, blocking local work traffic, also contributed to the increase of negativity during the pandemic (Jones and McGinlay 2020, 20). The most frequented places across the years can be found in the landscape around Edale and Kinder Scout, which are part of the 'Moors for the Future' initiative. The interactive webmap[28] that project provides on its website shows areas included in the peat restoration programme and areas heavily impacted by the moor

[28] https://www.moorsforthefuture.org.uk/our-work/moors-for-the-future-partnership-map

fires. That project also collected and archived the local history and stories from farmers, residents and ramblers between 2010 and 2012; these stories are a snapshot in time documenting important insider knowledge. That project provides a method for a rolling social media observation of trends and insights, which would allow reacting in real-time to the challenges. The hotspot areas of increased visitor influx match areas of peatbog restoration of the 'most degraded landscape in Europe'[29].

3.7 Conclusion and Further Research

As shown, Twitter data can be used to gain insight into trends and sentiments of the public based on specific key search terms. The interactive map of the most frequently mentioned locations in tweets follows the tradition of Stanley Milgram's mental map of Paris (Milgram and Jodelet 1992, 96, Fig.8.6) and Kevin Lynch's sketches of social urban landscapes (Lynch 1960). While these projects mapped social behaviour in urban landscapes, the same principle can be applied in rural landscapes such as the Peak District National Park, as shown in this study. The innovative method presented here offers an efficient and effective tool for location-specific entity recognition with an accuracy of almost 88%. This process can be repeated, scaled, and applied to other areas and landscapes. During the validation process, some issues of the automated process were identified by direct manual referencing of a sample dataset. Such challenges are inherent in automated analysis of natural language and the subject of the field of Natural Language Processing (NLP) and Named Entity Recognition (NER) (Ritter and Clark 2011). Further developments in this area, including advanced machine learning, could improve the process. Potential development of the method could lie in training a model that addresses the shortcomings of this algorithm and further development of the compilation of area-specific, sufficiently fine-grained location gazetteers.

---

[29] https://www.moorsforthefuture.org.uk/about-us

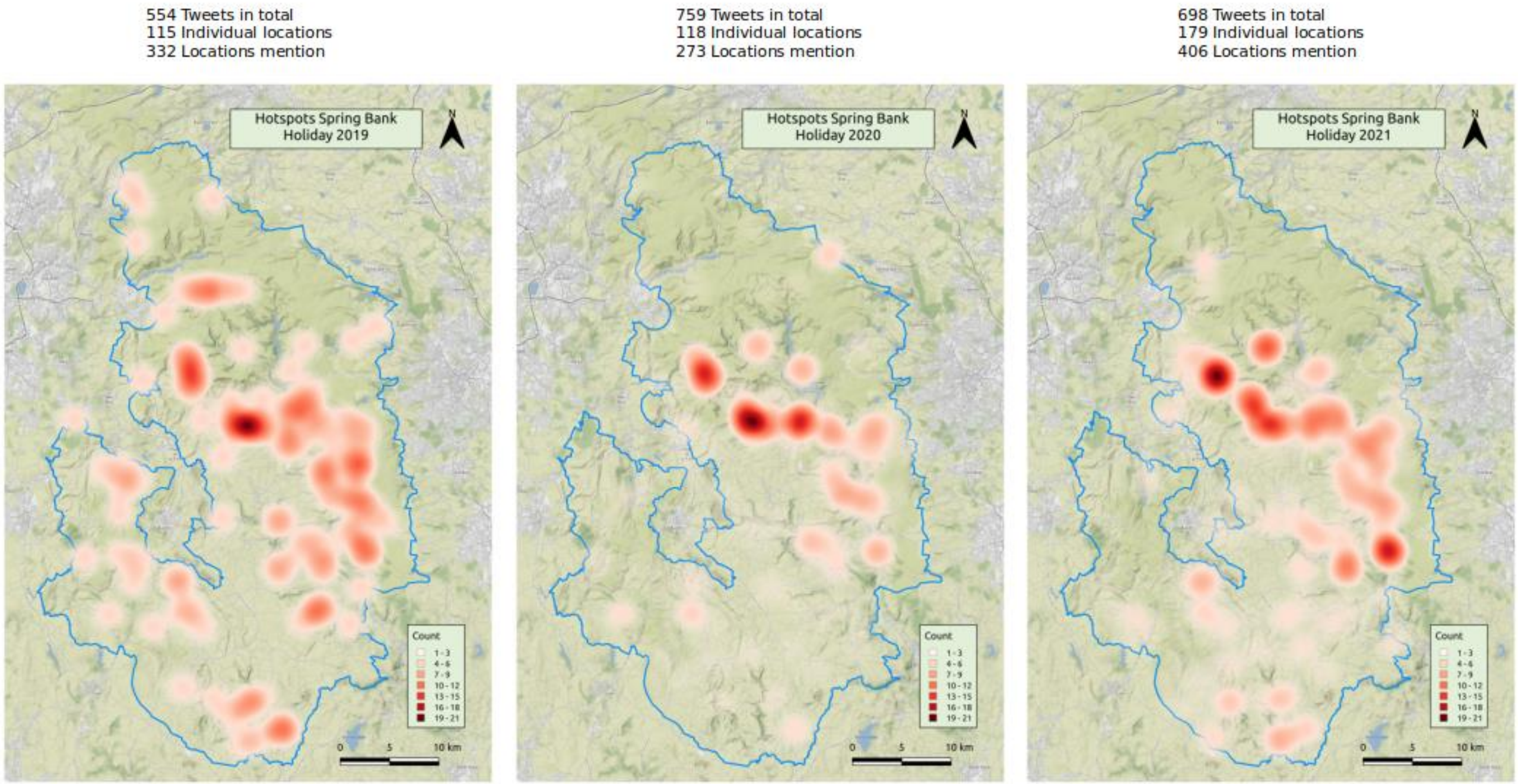

Figure 3-13: Heatmap visualisation of the years 2019 to 2021 showing the shift of locations mentioned in tweets visualised in QGIS (Map created by M. Tenzer, basemap Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL).

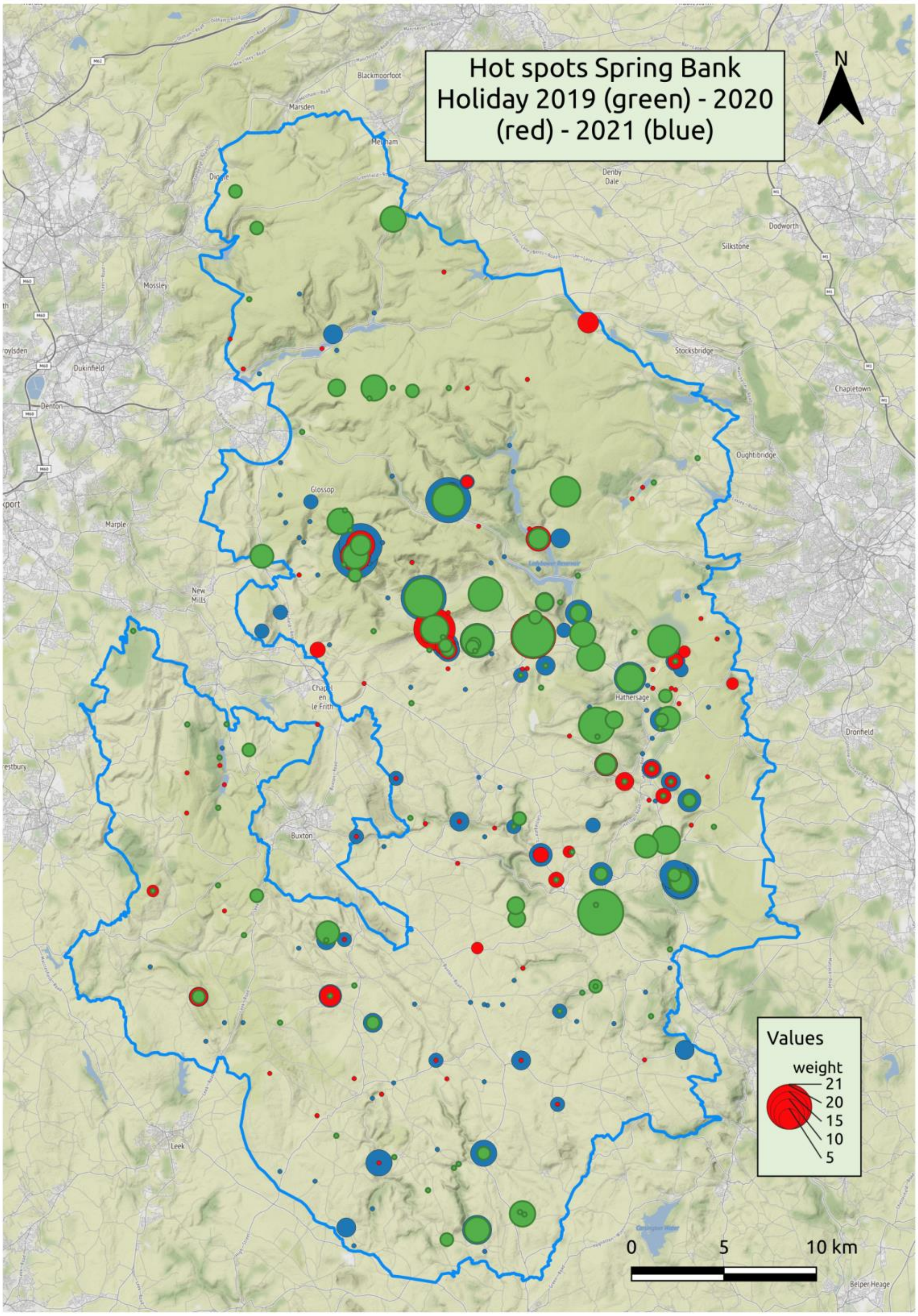

Figure 3-14: Another option for visualising hot spot locations mentioned in tweets for the years 2019-2021 in one 2D map: point size based on weight (frequency) visualised in QGIS (Map created by M. Tenzer, basemap Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL).

Extracting and visualising geospatial data from unstructured texts will generate data to benefit the public in three ways:

1. It will support heritage managers and local authorities managing places impacted by tourism and provide a background for planning and decision-making. Identifying locations with increased footfall, trending positively or negatively in public opinion, and gauging the emotional relationship to landscapes, monuments, and historic buildings will generate a more widely accepted and open planning process to facilitate change.
2. Further development of qualitative data analysis in research will provide methods and tools for interdisciplinary projects. Integrating data from unstructured text into the mapping process has so far led to the development of everyday applications, such as Google maps traffic alert and Google maps commute estimation.
3. The output of qualitative data analysis can produce visually attractive output formats, such as emoji clouds or heat maps, which convey complex information in a comprehensive and appealing form. This format will provide a broader basis on which communication between local authorities and communities can be based and encourage public engagement in change and development in extraordinary 'everyday landscapes’.

Cultural landscapes, such as the Peak District National Park, bridge the divide between natural and cultural resources. However, this rich and diverse environment also comes with challenges. To be better equipped to react to these challenges, this methodology and the results will be provided to the PDNPA to explore its application and utility in practice. The data will also be provided to the 'Moors for the Future’ initiative and the National Trust. The results of this project may provide background information for the management and protection of these precious resources while at the same time allowing people to enjoy the beautiful and rich landscape of the national park.

Acknowledgements

Coding support was provided by Dr Torsten Schenkel, Associate Professor at Sheffield Hallam University, Department of Engineering and Mathematics. I am also grateful for valuable discussions on analysis strategies and initial reviews of the draft versions of this article.

I would also like to thank my academic supervisors, Prof. John Schofield for initiating the idea of working with social media data and Prof. Julian D. Richards for his advice on best publishing and archiving strategies.

This project was carried out as a part of an AHRC/UKRI WRoCAH-funded PhD project at the University of York. Grant reference number: AH/R012733/1.

Supplementary material and additional information

The supplementary material for this paper consists of the code for the hashtag search, which was used to create the dataset for this research. It can be found in the **A. Appendix for Chapter 3: Appendix 1**. Full query code[30] and renv dependency file[31] are also available (**A. Appendix for Chapter 3: Appendix 2**). The code for the place specific NER is provided in the **A. Appendix for Chapter 3: Appendix 3**.

Furthermore, research using Twitter (now X) is currently not possible in the form as at the time this research has been conducted because the API is not accessible anymore. This is a common feature of the fast-evolving environment in which digital and AI research is situated. A further example for such a development is the interactive webmap, which was created and provided with the online format of the publication. The incorporated basemap provided by Stamen[32], a project funded by the Knight Foundation for the creation and publication of maps under the Creative Commons Attribution, has since been moved to be hosted by Stadia[33]. The Stamen

[30] https://intarch.ac.uk/journal/issue59/6/full-text.html

[31] https://intarch.ac.uk/journal/issue59/6/renv.lock

[32] https://maps.stamen.com/#watercolor/12/37.7706/-122.3782

[33] https://maps.stamen.com/stadia-partnership/

basemap is in several map tiles not provided anymore and, therefore, only features blank areas. However, changing the basemap in the webmap file provided for the publication cannot be amended. The map can be viewed in the Open Street Map[34] also integrated in the interactive webmap within the publication and provides, as a community driven map project more stable dataset.

[34] https://www.openstreetmap.org/about

Chapter 4:

Using Topic Modelling to Reassess Heritage Values from a People-centred Perspective: Applications from the North of England

Published as:

Tenzer, M., & Schofield, J. (2024). Using Topic Modelling to Reassess Heritage Values from a People-Centred Perspective: Applications from the North of England. *Cambridge Archaeological Journal*, 34 (1), 147-168.

DOI: 10.1017/S0959774323000203

Abstract

The historic environment—comprising a palimpsest of landscapes, buildings and objects—carries meaning and plays a crucial role in giving people a sense of place, identity and belonging. It represents a repository of ever-accumulating collective and individually held values—shared perceptions, experiences, life histories, beliefs and traditions. These social or private values are mostly ascribed by people to familiar places within this environment based on the ontological security which this everyday heritage provides. However, these values are notoriously hard to capture and categorize. This makes it difficult to incorporate them into heritage-management strategies, which typically rely on objective, fact-based datasets. In this paper, we present a new methodology to capture those elusive values, by combining Topic Modelling with the principles of Grounded Theory. Results show that our novel approach is viable and replicable and that these important values can be effectively and meaningfully integrated, thus creating more inclusive approaches to heritage management than exist currently.

4.1 Introduction

The historic environment is a generic and inclusive term that encapsulates the landscape and all places within it that are considered to be heritage. The historic environment is therefore a palimpsest, constantly evolving through the dual processes of change and creation (Bradley et al. 2004). As such, this environment carries meaning and plays a crucial role in developing a sense of place, identity and belonging for its occupants (Avrami, Mason and de la Torre 2000; Council of Europe 2000; 2005; Pendlebury and Gibson 2016; Ireland and Schofield 2015; Jones 2017; Pearson 1995; Schofield 2014; Stephenson 2008; Tuan 1990; West 2010). Furthermore, it is a repository of collective and individually held values. These values create communities ('heritage communities', after Council of Europe 2005) of shared perceptions, experiences, life histories, beliefs and traditions. Values are therefore routinely ascribed by people to their familiar landscapes, neighbourhoods and places based on the ontological security of this everyday heritage in daily life and routine (Grenville 2007). People and the landscapes they occupy are thus intimately enmeshed and meaningful places are deeply embedded in people's psyche, as demonstrated in recent research by Gatersleben et al. (2020), who used MRI scans to identify the activation of brain areas associated with emotional responses to such meaningful places in ways not found for either meaningful objects or neutral places.

People create a wealth of local knowledge and expertise through their everyday social life. This information is vital for understanding what makes places meaningful and valued. It should also be crucial for the management of those places to acknowledge the varied forms of perception vital for understanding what makes a place important to people. However, these collective and individually held values are diverse and hard to capture, making them difficult to incorporate into heritage management strategies that are historically designed around objective factual data[1],

[1] We focus on the presentation of the new method in this paper. A comparison with other qualitative approaches lies beyond the scope of this study (for this, see Jones 2017; Jones and Leech 2015; Nardi 2014).

typically based on datasets, such as the Historic Environment Records, finds databases and historic maps. The shortcomings we describe in the following section provide the justification for our approach.

4.2 Reviewing heritage values

A paradigm shift since the 1960s—the 'cultural turn'—had far-reaching implications for understanding the importance of values in the heritage sector and their creation through the meaning-making of communities and individuals (Tuan 1980). The nature of values was no longer seen as intrinsic in the fabric of the material world: buildings, landscapes, sites, or things (Pearson 1995, 308), but as a construct and the result of negotiation within democratic societies, communities and individuals, 'ascribed to heritage by society at large' (Avrami 2009, 179; Jones 2017, 21). Views on the role of heritage professionals and non-experts, and the concept of values in general, changed significantly—the heritage expert no longer had the monopoly of authority in defining heritage significance (Bonnell and Hunt 1999; Cosgrove 2004, 57; Pendlebury and Gibson 2016, 1-2).

However, while there is a consensus on the importance of recognizing people's perceptions and local knowledge for a more sustainable way to manage and think about heritage, practical solutions for the integration of people's voices in the decision-making process have been slow to develop and become integrated into heritage management. Heritage values are a particularly good example of these shortcomings. Ascribed values, when identified and determined by heritage specialists, can help to assess and determine the measures of management, inform statutory heritage protection decision making and provide support appropriate for specific parts of the historic environment. However, value categorization has been in dispute ever since the recognition of the benefits of positioning locally held viewpoints on value and significance alongside those of specialists, not least in terms of the tension created between the benefits of such an approach (e.g. its inclusivity) but also the significant challenges it entails (e.g. its practicality).

The suggestion to incorporate locally held values into heritage management systems was first promoted for heritage managers in the Amsterdam Charter (ICOMOS

1975), reinforcing the existing notion of public consultation and inclusion in heritage decision-making. The Burra Charter (ICOMOS 1979) did not mention this point before the 1999 version (now ICOMOS 2013). However, Smith's (2006) analysis of international conventions and charters, such as the Athens and Venice Charters, but also the Burra Charter, showed continued adherence to the principles of the Authorized Heritage Discourse (AHD), defined by Smith (2006) as a dominant concept of heritage management that is 'reducing the authority of non-experts' and 'appeals to the moral authority of expertise'. While the Burra Charter contains some progressive ideas about inclusion, its authoritative language still nonetheless strengthens the position of experts, albeit indirectly.

Public perception and participation was then integrated in the European Landscape or Florence Convention (Council of Europe 2000), the Faro Convention on the Value of Cultural Heritage for Society (Council of Europe 2005), and in Historic England's *Conservation Principles* document (English Heritage 2008). All these policy or guidance documents called into question the comprehensiveness and inclusivity of using only those objective, scientific value categories which provide the current framework (**B. Appendices for Chapter 4: Supplementary Material 1**). They challenged heritage managers to consider how much local knowledge and expertise is included in heritage evaluations and how fluid they are, to take account of the changing social fabric of an area over time, or people's changing opinions. Alongside Historic England, another national agency—Natural England—considers the natural and cultural value of landscapes based on the perception of people, aligning closely with the *European Landscape Convention*, referred to earlier. However, while current studies use 'proxies' or 'secondary measures' to identify people's perception (for example, tourist numbers, numbers of footpaths and car parks as 'opportunities for health walks', or 'measures of accessibility': (Natural England 2015, 10), no time-efficient and effective method for the collection of qualitative data has been developed, mirroring the situation for cultural heritage[2].

[2] They are also recognizing how problematic it is to distinguish the two (Harrison 2015).

With these developments in mind, we ask: is the current value system on which experts assess heritage fit for purpose? And is the valuation framework for conservation, arguably exemplified by that currently offered by Historic England (English Heritage 2008), sufficiently forward-thinking and up to the task in a fast-changing world? This question of changing perceptions aligns with the growing realization that the term 'heritage' itself is fluid and dynamic: a social process of creating meaning and memory, communities and identities. Value categories, it was argued, should also be sufficiently flexible to adapt to societal change and environmental pressure (Byrne 2008b; Dalglish and Leslie 2016; Harrison 2010; Jones 2017; Jones and Leech 2015; Waterton, Smith and Campbell 2006).

Against this background, we explore and critique the current value categories described in *Conservation Principles*. As a significant departure from current applications, we propose an innovative approach which involves identifying those hard-to-obtain but important values held by local people and their communities. This approach will address the lack of inclusion of people's perceptions and the dynamic quality of social values[3] as an ongoing challenge for heritage and landscape management[4]. We use Topic Modelling to identify latent or emerging value themes

---

[3] To avoid a confusion of the widely used term 'social value' as identified by the Burra Charter (also after Johnston 1992), we decided to avoid this term as a subcategory for the Communal category as defined by English Heritage (2008). The category has been renamed as 'private value' including the family history and personal connections of individuals (see also Modesto and Waterton 2020).

[4] It has to be noted that the data source and method of data collection represent a convenience sample and are limited and biased by various factors inherent in this approach, i.e. online accessibility, IT literacy, general interest in heritage, interest in participating in local heritage, exclusion from access to the countryside and heritage, etc. The survey showed that 54.2 per cent of participants were in the age brackets of 56–75 and a further 20 per cent in the age bracket of 46–55. 91.6 per cent identified themselves as White British which reveals a limitation of the result towards the dominant resident group. 54 per cent were female participants as opposed to 43 per cent male, 1 per cent non-binary and 2 per cent who answered 'preferred not to say'.

from a public survey, providing the individual stories (the intangible element) of 'everyday heritage'. This offers an open-minded approach to qualitative data. We propose Natural Language Processing (NLP) and Machine Learning (ML) with a focus on Topic Modelling as a method, following basic principles of Grounded Theory for improving the system of valuation in heritage management (Charmaz 2006; Creswell 2017, 276). Following Grounded Theory means a first investigation of the data free of the researcher bias[5]—with no predefined codes or assumptions. The themes latent in the data are explored as they emerge and in a later phase of the study are structured into a framework of topics. The results are correlated with the current framework of values as applied by Historic England, suggesting how new categories could help to address the changing expectations, needs and demands of the public, partly developed as a reaction to a changing world after the COVID-19 pandemic and the increasing pressures on natural and cultural landscapes that have emerged as a result (Ginzarly and Srour 2021; Historic England 2022b; UNESCO 2021b).

4.3 Data sources and methodology

4.3.1 Overview of the method

To explore ways to capture social values in a form that can be used within the planning process and in heritage management decision-making requires a method that is efficient and scalable. Here we use Topic Modelling to present place attachment in a

[5] As a researcher bias, we define the preconceived assumptions that are noted at the beginning of a qualitative research process to ensure that this bias is clarified throughout the process. However, other sources of bias are introduced in other stages of the research similar to other qualitative methodologies.

format that can be used for categorizing individually held values in cultural landscapes. We will first describe the method before going on to deliver a proof of concept. While this novel methodology is based around two areas in the UK, it is transferable to

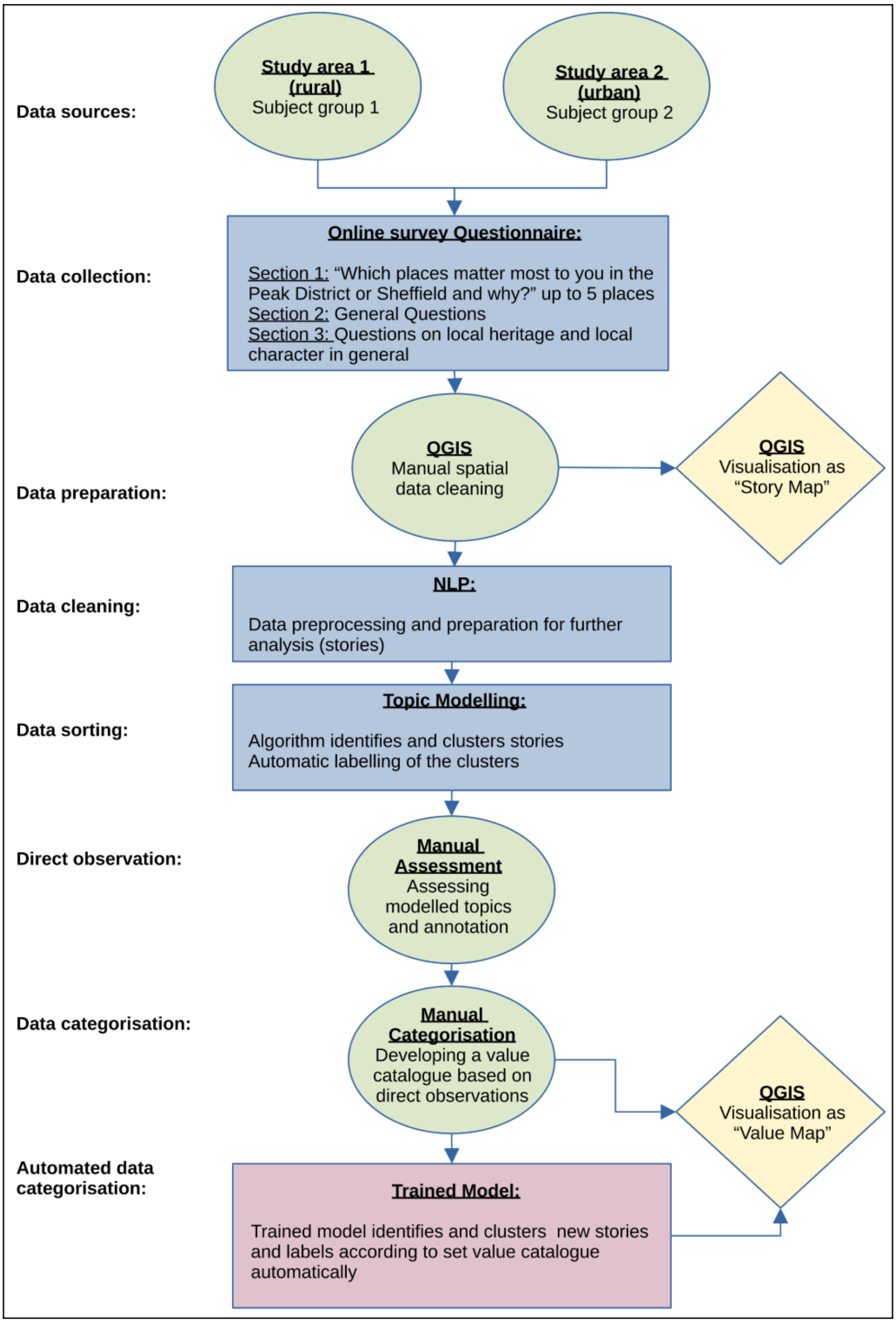

Figure 4-1: Detailed methodology developed in this research. The aim is to create the topics based on the 'stories' of survey participants and correlate these with value categories as set out in Historic England's Conservation Principles (English Heritage 2008), extending the current value catalogue. (Green: manual process; blue: computed process; red: outlook; yellow: visualization.)

any context where local authorities have capacity to routinely use survey data or online portals to record people's interactions with or feelings about a place.

The flowchart in Figure 4-1 details the process applied in this research. Residents of two study areas were asked to provide five favourite places that matter most to them and ‘stories’ about their personal connection to these places. The stories were then fed into a Topic Modelling algorithm to get a first, bias-reduced insight into the latent themes within the data and preliminary topic labels summarizing the topic clusters. Direct observation then allowed evaluation of the modelling result and an assessment of the usefulness of this approach. The annotated data were subsequently categorized based on the value catalogue provided by Historic England's *Conservation Principles*. Visualization in GIS can be created at two stages of this process: in the form of a ‘story map’ as the basis for communication between local authorities and residents; and in the form of a ‘value map’ following the process of categorization, with the potential to provide background information for planning purposes.

4.3.2 Study areas

Two study areas were identified for this project: the City of Sheffield and the Peak District National Park, both located in the north of England (Figure 4-2). These locations were chosen to offer insight into the enmeshed relationship between people and places amongst a combination of urban and rural communities. Additional methods include in-depth interviews and a social media analysis (Tenzer 2022; Tenzer and Schofield 2023). While the district of Sheffield overlaps with the National Park, the two landscapes are in many respects different and distinctive. Residents of both areas are closely connected to the landscapes of both study areas through leisure activities, work and shopping, cultural activities and educational organizations. This integration was considered an advantage when selecting the two study areas.

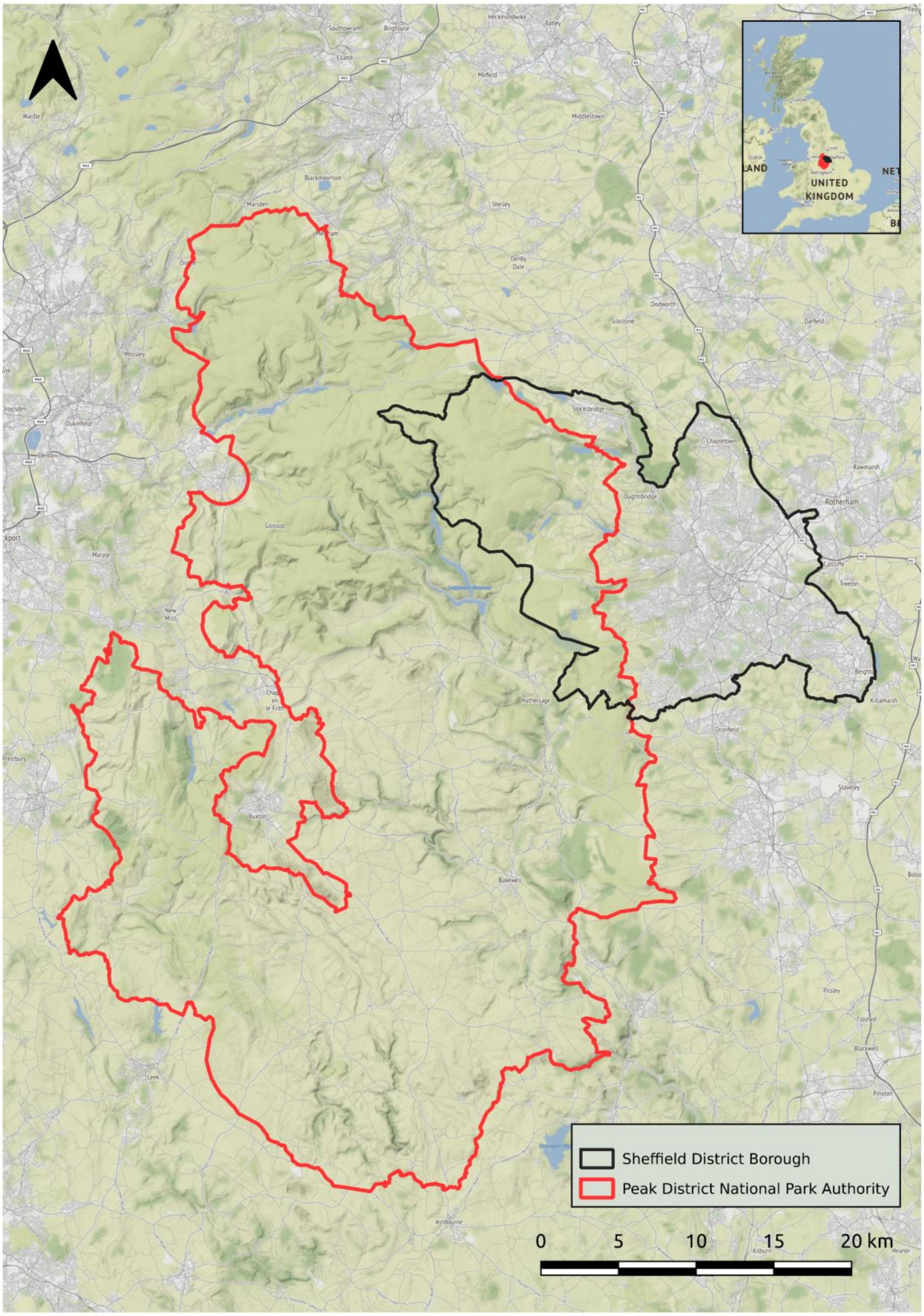

Figure 4-2: Study areas: Peak District National Park and Sheffield city. Residents of these areas were invited to participate in a survey and provide up to five favourite places within the study areas and their 'stories' of personal connection. (Map created in QGIS; data contain OS data © Crown copyright and database right 2022. Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under OdbL.).

The Peak District National Park—the first designated National Park—covers an area of 1438 sq. km at the time of writing (PDNPA n.d.). The Park includes various archaeological and historical sites from the Palaeolithic to recent periods, including prehistoric burial mounds, henges and stone circles, medieval field systems and settlements and post-medieval industrial sites, including stone quarries. It is a place of residence for approximately 38,000 people. At the point of writing, the Peak District National Park has 2900 listed buildings, 109 conservation areas and 450 scheduled monuments, all of which involve various degrees of statutory protection (PDNPA n.d.).

Sheffield lies to the west of the Peak District National Park. It is dominated by seven hills and two universities, covering an area of 367 sq. km with approximately 556,500 residents. It is best known for its industrial heritage. Today, both visitors and residents favour the town for its multicultural character, the wide offer of cultural events and the quality of an outdoor city with its vicinity to the National Park. There are currently 1200 listed buildings, 38 conservation areas and 43 scheduled monuments in the city[6].

4.3.3 Survey method

Residents of the Peak District National Park (PDNP) and the City of Sheffield were first invited to participate in an online survey. The online questionnaire was published through the channels of both local authorities, the Peak District National Park Authority (PDNPA) and Sheffield City Council, comprising their websites, social media channels (i.e., Facebook, Twitter, Instagram, LinkedIn) and a specific mailing list of over 4500 recipients in Sheffield. Participants therefore included residents of the two study areas exclusively[7], who self-identified as residents of either study area. In total 476 responses were received. Forty-eight participants identified themselves as residents of the PDNP and 386 as residents of Sheffield; 42 participants did not answer this question.

---

[6] https://www.sheffield.gov.uk/

[7] This was reinforced in the publication text on social media and as one of the first survey questions (see **B. Appendices for Chapter 4: Supplementary Material 3**).

The survey questionnaire consisted of three sections: general information (such as place of residence, age, ethnicity, and level of education); the question which places matter most to the respondent, and why; and questions on their perception of the local heritage and landscapes (see **B. Appendices for Chapter 4: Supplementary Material 3**). Questions in sections 1 and 3 were multiple-choice. To allow the most flexibility and subjectivity in the second section, respondents were given the opportunity to enter a location on a map and their 'story'—the personal reason for a connection to the place—as a free-text entry of up to 300 words. To identify more places of individual importance, the respondents were asked to enter up to five places in one questionnaire. Not all respondents provided five places as requested. The average was just over one place/response per respondent, resulting in 547 places, of which 509 provided a story, experience or perception of heritage, landscapes or history relating to one of the two study areas.

The survey responses were downloaded and imported into a Geographic Information System (GIS)[8]. The dataset was then cleaned. This included: eliminating places not located within the study areas and repositioning locations set in the wrong place (determined by written input in the location field). The participants provided 174 favourite places in the PDNP and 298 in Sheffield. The intersection of the study areas included 60 locations. Eleven locations close to but outside the study area limits to the east of the City of Sheffield were included as the locations were close to the boundary, i.e., Buxton (see Figure 4-3).

The online questionnaire for this research followed the principle of offering a low-cost/cost-free, practical solution for survey design by using the Qualtrics software[9] and an embedded Google Maps map. This questionnaire allowed the participants to use a familiar map interface to locate and pin a location, which automatically provided geospatial coordinates for the GIS map analysis to create georeferenced stories. Where participants were not able to locate the place on the map, they could also enter locations as free text. In this case a Named Entity Recognition process was used as

---

[8] In this project we use QGIS, a free, open-source platform https://www.qgis.org/en/site/

[9] https://www.qualtrics.com/blog/qualtrics-survey-software-free/

detailed in Tenzer (2022). The data were then ready to be preprocessed for analysis with NLP and TM.

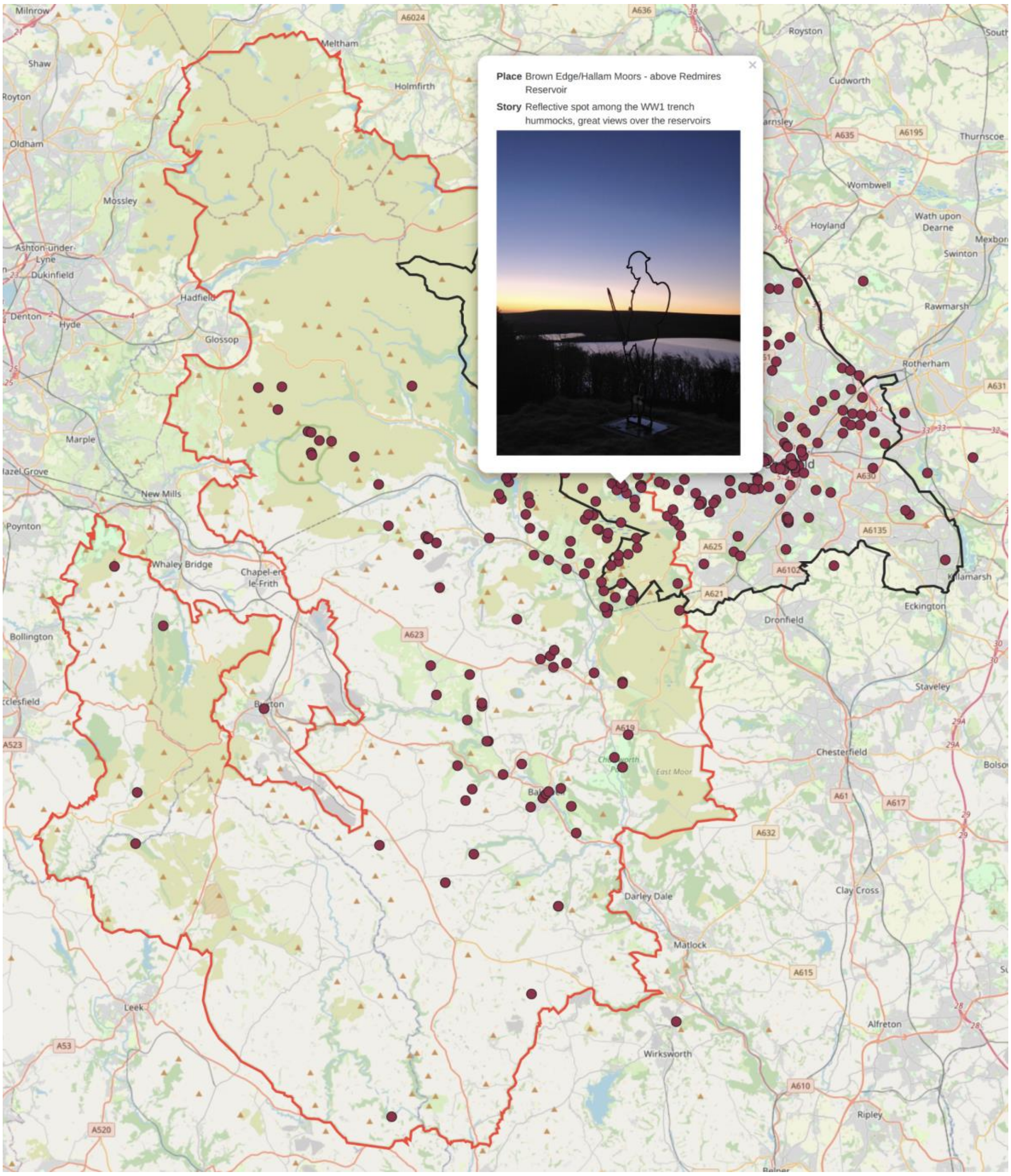

Figure 4-3: Interactive 'story map', showing the favourite places provided by residents of the two study areas who participated in an online survey. The participants were asked to provide up to five places that matter most to them and give an up to 300-word description of the reasons for the connection to these places. (Map created in QGIS/Leaflet, © OpenStreetMap contributors; data contain OS data © Crown copyright and database right 2022.)

4.3.4 Natural Language Processing and Topic Modelling

NLP and TM are emerging methods for data analysis that are particularly relevant to qualitative research. Heritage and landscape studies have yet fully to appreciate and integrate the opportunities offered by these tools (but see Bordoni, Mele and Sorgente 2016; Condorelli et al. 2020; Fiorucci et al. 2020; Matrone et al. 2020; Verschoof-van der Vaart et al. 2020). The Council of Europe actively encourages the use of Artificial Intelligence (AI) in all sectors, including heritage (Traviglia 2022), as a result of increasing interest in Text Mining and Machine Learning/Topic Modelling for computational language analysis (Goerz and Scholz 2010; Sassolini and Cinini 2010; Sporleder 2010). Our research applies AI in line with these projects, using NLP algorithms, to model categories of values based on perceptions, experiences, concerns and visions of residents of the two study areas, as described in the 'stories' provided by the online surveys.

TM is a time-efficient method to analyse qualitative data and has been tested for its capabilities elsewhere (e.g. Abram, Mancini and Parker 2020; Cai et al. 2021; Franzosi, Dong and Dong 2022; Ginzarly and Srour 2021). More importantly, this method of text analysis allows themes and topics emerging from or 'latent' within the data to be captured without preconceived codes. This approach aligns with the underpinning elements of Grounded Theory (e.g., Charmaz 2006; Odacioglu and Zhang 2022). Grounded Theory is usually applied 'when little is known about the phenomenon' (Chun Tie, Birks and Francis 2019), or in the case of this research, where traditional thinking dominated by experts (often referred to in the context of an Authorized Heritage Discourse: see (Schofield 2014; Smith 2006; Waterton, Smith and Campbell 2006) is preferably avoided, exploring the connection between people and places, perception and place-making, in a narrow and individual sense, without the bias of preconceived expert knowledge.

The following steps are part of the analysis shown in the flowchart in Figure 4-1. To undertake this analysis, the survey data were downloaded. Coordinates, place names, stories and photographs were extracted from the survey dataset—separating these from the demographic data—and then cleaned and preprocessed with *textmineR*, eliminating places not located within the study area and relocating pins set in error. The resulting data set was then fed into the NLP toolchain using R and Python routines in RStudio.

The five steps of the Topic Modelling process are as follows:

1. Data wrangling: cleaning text, lemmatizing, stemming, removing stop words, creating a DocumentTermMatrix.

2. Choosing a model algorithm: the R package textmineR provides an innovative method of topic modelling and labelling (currently under development at George Mason University, USA, by Tommy Jones: see Jones 2021; Jones, Doane and Attbom 2021). This unsupervised machine-learning model uses unlabelled data and implements the Latent Dirichlet Allocation algorithm and Gibbs Sampling (Blei, Ng and Jordan 2003). It creates word clusters and identifies topic labels based on the probabilistic distribution of words over topics and topics over documents. This means that each topic is a combination of keywords which have a specific importance (weight) contributing to the weightage of each topic.

3. Running the model and adjusting parameters: most of the parameters were left to default settings, while the number of topics and iterations was optimized. The model iterates n-times over the data and provides information on the topic coherence level at k topics (Figure 4-4). This means that the model in our case attempts 4000 times to make more sense of the relationship between the documents and to create more meaningful clusters where documents relate more closely to each other.

4. Choosing the optimal number of topics: coherence (Figure 4-4) measures the degree to which the documents (stories) in a topic show high semantic similarities and support each other in their statements. This gives us the model we want to choose for further analysis. In our study, 40 models were created, and the best coherence is provided at 35 topics (tqx94 2022). After 35 topics, the coherence score flattens

out, meaning that a greater subdivision of the data would not provide any more coherent and meaningful topic groups. The coherence is not very high in our case and more data would help to fine-tune the model. Nonetheless, this gives a good starting point for the manual analysis.

5. Topic labelling: labels are based on the most frequently appearing key terms in the documents/stories of each cluster (bi-grams: two closely associated words in a text) (Figure 4-5)(tqx94 2022). Labels created in this process are not meaningful titles, but give a good indication of the dominant theme in each cluster.

Using this unsupervised learning method prevents the introduction of researcher bias in the next step of data analysis, which is comparable to the coding phase of text in other approaches, such as NVivo, as the algorithm has completed the clustering independently and created a pre-labelled dataset. This approach follows to the tenets of Grounded Theory (Charmaz 2006). Our attempt to capture the themes latent in or emerging from the data themselves opposes the commonly used and widely criticized approach of applying codes or categories and trying to fit the data into these codes, introducing bias and issues with reliability and validity (Banks et al. 2018; Welsh 2002). The approach of creating two bi-gram label variants per topic has proved to help identify the themes more accurately (Label 1 and Label 2) (Table 4-1).

Following the Topic Modelling process, the dataset was exported and the labelling by the algorithm was manually assessed in a direct approach to observe the coherence of the topics (Lau, Newman and Baldwin 2014). In the same step, the most dominant topics were identified manually and compared to the suggestions made by the algorithm.

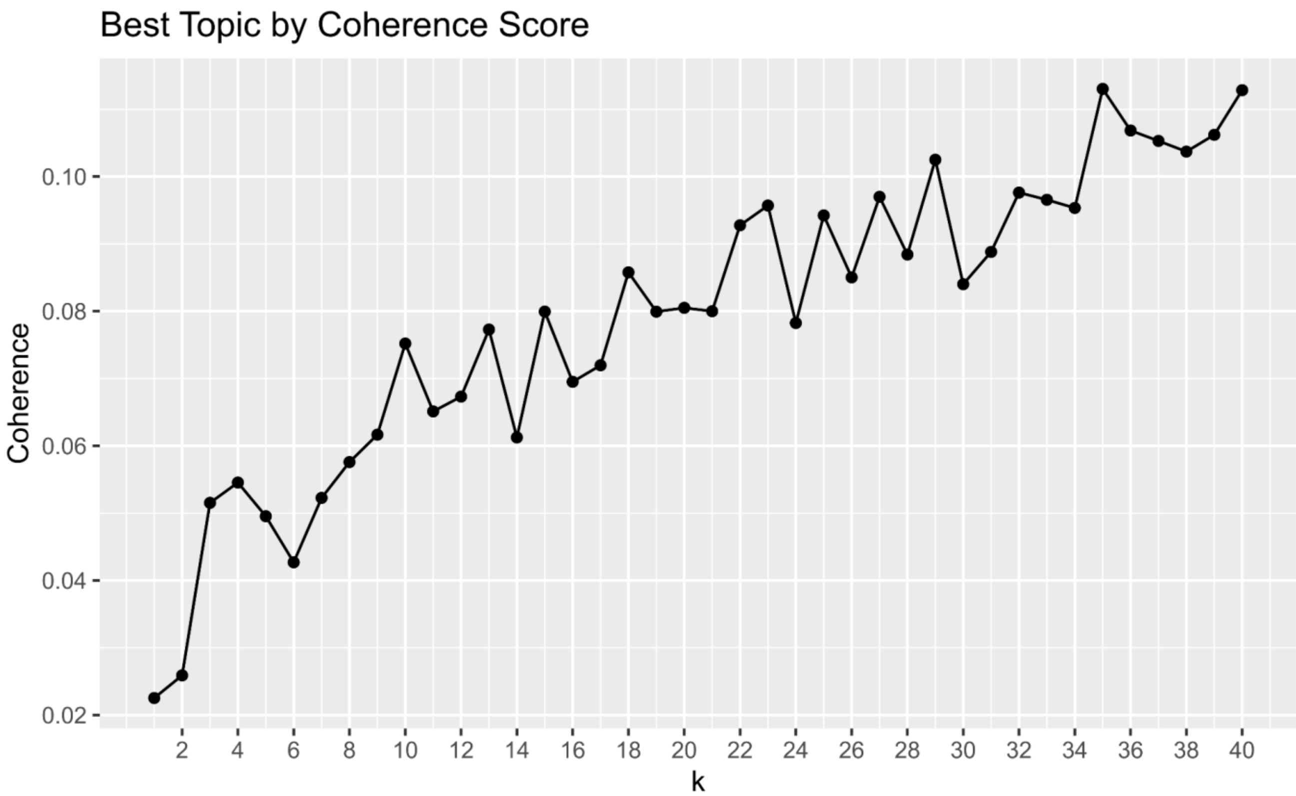

Figure 4-4: The Topic coherence gives the degree to which documents in one cluster are closely related. This depends on the size of the cluster and the number of topics chosen. For example, if we subdivide the documents into four different topics, the documents have less coherence—similar semantics and supporting the statements of each other—than at the point of 35 clusters, where the number of documents in one cluster is smaller but the relationship between the documents higher. At 35 topics the coherence is highest, flattening out with more subdivisions. Therefore, 35 topics will provide the best first insight into the latent topics within the documents (tqx94 2022).

A manual evaluation has shown that some of the topics are related to heritage, history and archaeology (Robin_Hood, industrial_heritage, steel_industry, list_building), while others are related to private life and community (friend_family, happy_memory) or aesthetic values (great_view, natural_beauty). However, topic labelling still needs human input and dataset structuring for meaningful labelling and categorization of the data. The application of the algorithm does not fully replace manual assessment and fine-tuning of coding (Cai et al. 2021; Chang et al. 2009; Leeson et al. 2019). In our data model, the manual categorization was subsequently carried out to understand on what basis the documents were clustered (top terms) and if the clustering proved meaningful.

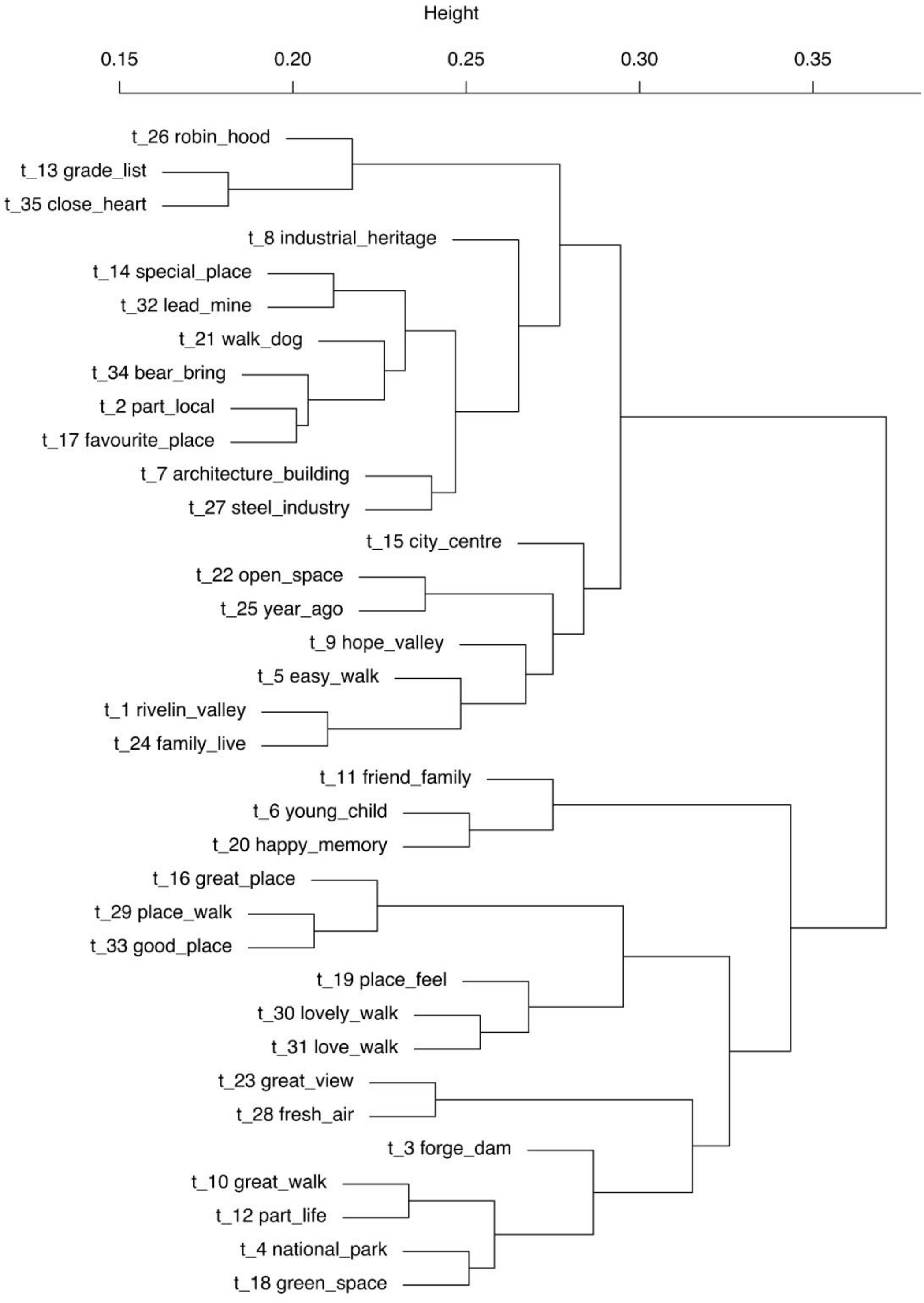

Figure 4-5: The Cluster Dendrogram shows the 35 topic clusters with their respective labels chosen by the modelling process. The levels (Height) show the similarities of topics. Topics in the lower level of the diagram on branches close together show topic cluster that have themes closely related to each other. For example, topic 10: 'great_walk' and topic 12: 'part_life' consist of documents or 'stories' with content that support each other and can be summarized under the respective label. Similarly, topic 4: 'national_park' and topic 18: 'green_space' consist of documents with a similar theme. At a higher level, all four clusters are related to each other—to a lesser degree than at a lower level but more closely related than, for example, topic 3: 'forge_dam', again, a level higher up.

Table 4-1: Following the Topic Modelling process, the optimal number of topics (35) was chosen to create labels for the topic clusters. The labels are based on the most frequent key terms in the documents/stories of each topic cluster. Labels vary in their quality, which makes a manual evaluation necessary.

Topic	Label 1	Label 2
1	rivelin_valley	bear_bring
2	part_local	start_point
3	forge_dam	play_area
4	national_park	car_park
5	easy_walk	endcliffe_park
6	young_child	child_grandchild
7	architecture_building	back_yard
8	industrial_heritage	industrial_history
9	hope_valley	walk_edge
10	great_walk	natural_beauty
11	friend_family	meet_friend
12	part_life	amaze_view
13	grade_list	list_build
14	special_place	post_office
15	city_centre	close_city
16	great_place	place_visit
17	favourite_place	enjoy_walk
18	green_space	rich_history
19	place_feel	walk_home
20	happy_memory	lot_memory
21	walk_dog	cricket_pitch
22	open_space	botanical_garden
23	great_view	easily_accessible
24	family_live	family_tree
25	year_ago	live_year
26	robin_hood	beautiful_build
27	steel_industry	portland_work
28	fresh_air	good_view
29	place_walk	good_place
30	lovely_walk	place_walk
31	love_walk	walk_area
32	lead_mine	geological_historical
33	good_place	huge_amount
34	bear_bring	bakewell_pudding
35	close_heart	area_close

4.3.5 Value categories and perception correlation

In the third step, the data were manually coded and the quality of the topic clusters and labels was assessed. To correlate these emerging values based on the experiences and perceptions of people with the value categories set by HE (Historic England since 2015, formerly English Heritage) (English Heritage 2008), each document was allocated to one of the following value categories (subcategories): evidential value, historical value (associative, illustrative), aesthetic value (design), communal value (social (renamed to 'private'[10]), commemorative, spiritual). The subcategories of 'Communal' values were developed to be more nuanced regarding the variety of aspects of stories, accommodating elements that did not fit into the HE categories. Also, a new category capturing the concept of nature in people's perception of landscape values was integrated to address the increasing awareness of valuing the environment in view of climate change and biodiversity loss (Table 4-2). This was also necessary to overcome

Table 4-2: Historic England's value categories as set out in the Conservation Principles (English Heritage 2008). Additional nuance to these categories is proposed based on the underlying themes identified in survey data. The additional values (shaded cells) address societal trends and the changing and dynamic demands and needs of residents in the Peak District National Park and the city of Sheffield following the COVID-19 pandemic.

Categories Historic England	Sub categories	Description
Evidential Value		Evidential value derives from the potential of a place to yield evidence about past human activity.
Historical Value	Illustrative, Associative	Historical value derives from the ways in which past people, events and aspects of life can be connected through a place to the present.
Aesthetic Value	Design	Aesthetic value derives from the ways in which people draw sensory and intellectual stimulation from a place.
Communal Value	Private	Private (compared to public) heritage values are commonly overlooked as "family history" with no place in the generalisation of heritage for the common good.
	Commemorative, Spiritual	Communal value derives from the meanings of a place for the people who relate to it, or for whom it figures in their collective experience or memory.
	Arts & Culture	Aspects of landscapes and environments that afford a communal experience of culture, entertainment or learning.
Green space Value	Health	Nature values derive from the quality of green spaces, biodiversity, wildlife. Health value derives from the qualities and opportunities of a place or landscape that provides space for outdoor activities, in particular developed during the COVID-19 pandemic.

[10] For 'private values' see [3].

the diametrical-opposite division developed in value systems over the past decades, dividing cultural/natural landscapes, tangible/intangible elements, and learning/mental health approaches to heritage—a subdivision rarely evident in the daily experience of the environment.

4.4 Results

In this section we present the results of the modelling process, direct observations and manual annotation for correlation with the values defined in Historic England's Conservation Principles. We do this by first introducing the results of the TM. We then describe the manual observation of usefulness of the modelled topics and labels. Finally, we summarize the results of the correlation and describe the development of a more nuanced division of communal values as identified in the survey data.

Results show that people's perceptions correlate with some of the value categories of HE, showing that the expert definition of heritage values is capable of capturing parts of the individually held values. Our bottom-up approach can be aligned with the expert-led approach to find a common ground for heritage value categorization.

4.4.1 Modelling categories

As suggested in Figure 4-4 and illustrated in Figure 4-5 and Figure 4-6 the modelling computed 35 clusters as an optimal topic number. The sizes of the clusters varied from five documents/stories (Topic 35) to 33 (Topic 6).

The TM approach provided mixed results. Some documents did not provide a story and were not included in the clustering (n = 37). Other documents provided stories that were too short (just one or a few words), which did not allow the algorithm to cluster in a meaningful way. These documents were labelled 0 (n = 20). However, during the manual process, these documents could still be allocated to one of the categories. For example, ‘magical place’ was allocated to the subcategory

'Spiritual' value[11]; the location 'Surprise View' with the description: 'In a word, "breathtaking"' containing the keyword 'view' in the location name was associated with 'Aesthetic' value; the location 'Kinder Ravine' provided the description 'Scrambling', which was annotated with the code 'outdoor activity' and associated

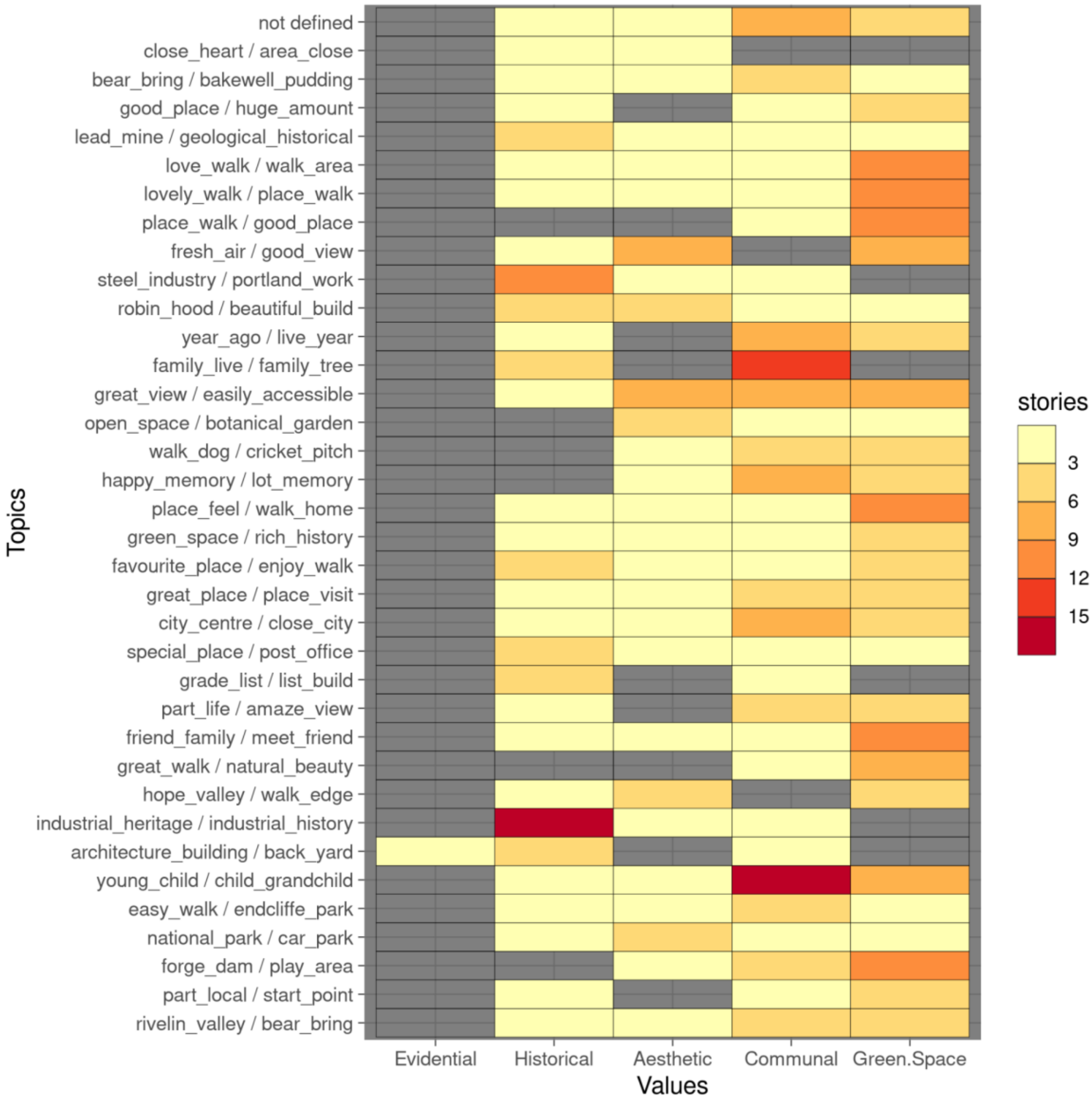

Figure 4-6: Result of manual evaluation of the topic modelling process. Document count over topics, showing the number of documents (stories) allocated to the corresponding value categories as defined by Historic England. High counts (red) show the dominant value category in a topic cluster. Rows with yellow cells (low counts) show incoherent topics with high variation in values. The proposed category for 'Green Space' value, including 'Health' value, dominates in the assessment.

---

[11] We recognize that this blunt interpretation could introduce interpretative limitations and that labelling a place described as 'magical' as 'spiritual' may not align directly with conventional definitions of spirituality. However, categorization demands a degree of flexibility and decisions that may not be free of limitations.

with the type ‘Green Space’ and the subcategory ‘Health’. In total, 510 documents provided sufficient textual data for the Topic Modelling algorithm to process.

The top-terms list created during the TM process clustered the most frequent words of each document, which provides the basis for the clustering decision and the labelling of each topic cluster. This list was used to decide on the most important keywords of each story or document. The resulting contracted list was subsequently condensed into a single one-word code. This code was used to allocate the document to one of the four value categories included and defined in Historic England's Conservation Principles (English Heritage 2008) (historical, evidential, aesthetic, communal[12]) (Table 4-2).

4.4.2 Value category development

Automated Topic Modelling provides valuable first insight into the latent themes in the survey data. The stories showed clear trends and commonalities (Table 4-1). However, it was necessary to assess and refine the categories manually. It became clear that some of the documents could not be fitted into this predefined set of values defined in Conservation Principles; for example, topics labelled with ‘calm_place’, ‘great_walk’, or ‘nature_reserve’. A heatmap shows that survey respondents based their place attachment predominantly on communal, green space and health aspects (Figure 4-7 and Figure 4-8) with green space and health dominating the relationship between people and places. A more nuanced approach to defining heritage values might extend the existing framework by further subdividing the category of communal values (private, spiritual, commemorative) to include ‘Arts & Culture’ (Figure 4-7; Table 4-2)

---

[12] ‘Communal value’ represents the shared value of communities, which is based on shared histories, beliefs, or myths. While these are established independently from the individual, an accumulation of individual values can form a different level of shared value when present in a wider group of the community and are, therefore, a different facet to the communal value, such as commemorative or spiritual.

Figure 4-7: The category of communal values can be subdivided into private, spiritual and commemorative values, as set out in Historic England's Conservation Principles. Themes emerging from the survey data show a trend to connect places based on green space, health and arts & culture. Yellow colours in the heatmap represent low numbers of stories, while darker red colours represent higher numbers of stories categorized in the respective categories. The graph shows that survey respondents based individual place attachment predominantly on aspects of communal, green space and health aspects.

to capture the intangible qualities of the cultural aspect of areas in Sheffield, in particular. Furthermore, a new (sub)category of 'Green Space' would help to overcome the artificial divide between cultural and natural heritage which people rarely recognize (Byrne and Ween 2015; Harrison 2020; Latour 1993). This category would capture the values people place on wildlife and ecology, for example, while addressing the increasing awareness of environmental pressures on natural resources. Another subcategory of 'Green Space' value could be defined as 'Health' value (both physical and mental), which encompasses, for example, people's COVID-19 experiences and the different approaches to landscapes that have emerged as a result of the lockdowns that were imposed at this time alongside relaxation and the calming qualities of woods in Sheffield's parks that relate to public benefit.

We can illustrate this argument with some examples. First, the document cluster of Topic 19 consisted of 18 documents. Label 1 was given as 'place_feel' and label 2 as 'walk_home'. A list of the most dominant words in the stories showed that 'lockdown' was mentioned in five documents. Ten documents were labelled with the broad subcategory 'Health'. These documents mentioned 'access' to 'nature' or 'green spaces', 'peace', 'walking', 'calming' and 'relax'. Three documents were allocated to 'Aesthetic' and 'Green Space'. The documents which were not allocated to the aforementioned categories were categorized as 'Aesthetic (Design)' (n = 3), 'Historic (Associative)' (n = 2), 'Natural Resort' (n = 2), and 'Arts & Culture' (n = 1). Topic cluster 24 was labelled by the algorithm as 'family_live' and 'family_tree'. Of the 18 documents in this cluster, 12 were manually categorized as relating to 'Private' value and two to 'Spiritual' value. This exemplifies another cluster of high coherence and close association with the modelled labels.

Cluster 27 comprised 12 documents labelled as 'steel_industry' and 'portland_works'; the manual assessment in this case showed that nine documents could be categorized as 'Historical' value, with the distinction of 'Associative' and 'Illustrative'.

Figure 4-6 shows the distribution of documents (stories) over the developed catalogue of heritage values, as the result of TM and manual assessment/refined categorization. The colour scale indicates the number of documents for a topic cluster within each cluster. Dark red cells indicate a high number of documents in one category and, therefore, a good correlation of document content in the Topic Modelling process. Rows with yellow cells show a great variety of value categories in one cluster and low coherence in the respective cluster overall. Furthermore, the graph shows the value categories most dominant in the perception of survey responses. 'Communal' values were mentioned 142 times, 'Historical' values 104 times, 'Aesthetic' values 79 times and 'Green Space' qualities 175 times (of these, 119 documents were categorized as ‘Health’). Evidential value could only be identified in one document based on the description of the connection to a place. The topic allocation Figure 4-5 can be correlated with Table 4-2. For example, Figure 4-6 shows Topic 8 to be most dominant, categorized as ‘Historical’ with labels in Table 4-2 defined as ‘industrial_heritage’ and ‘industrial_history’. This example shows a case where manual annotation confirmed the outcome of the topic modelling process. The dominant topics and categories in the proposed categories ‘Cultural’ value and the subdivision of the introduced category of ‘Green Space’ values with ‘Health’ can be seen in Figure 4-7. Apart from the category ‘Communal’ values, ‘Green Space’ and ‘Health’ feature most strongly in public perceptions.

In general, topics labelled as ‘young_child’ (Topic 6), ‘friend_family’ (Topic 11), ‘happy_memory’ (Topic 20), ‘family_live’ (Topic 24), ‘close_heart’ (Topic 35) were most dominantly associated with ‘Communal’ values (Spiritual, Private), while labels such as ‘place_walk’ (Topic 30), ‘love_walk’ (Topic 31), ‘place_feel’ (Topic 19), were associated with ‘Health’. Documents of the category ‘Historical’ value were labelled ‘lead_mine’ (Topic 32), ‘steel_industry’ (Topic 27), ‘grade_list’ (Topic 13), ‘industrial_heritage’ (Topic 8), ‘architecture_building’ (Topic 7).

Examples of stories provided by survey participants which qualified for the different categories are provided in Supplementary Material 2[13]. For example, category ‘Green Space’ with the subcategory ‘Health’ frequently referred to ‘lockdown’ and the COVID-19 pandemic. Unsurprisingly, these areas were primarily located in the city of Sheffield, where people had to find open spaces for appropriate social distancing. A theme detected and correctly labelled by the algorithm was the practice of ‘ash scatters’ allocated to ‘Communal’—more precisely ‘Spiritual’ value—primarily associated with specific open-space and viewpoints in the Peak District National Park. Another subcategory of ‘Communal’ value is shown here as subcategory ‘Arts & Culture’.

These examples show that Topic Modelling found coherent topic clusters and labels suitable for a first insight and open-minded approach to the data with no

[13] Available in **B. Appendices for Chapter 3**.

preconceived assumptions and codes. The clustering can be useful for a first analysis as a basis for the preliminary coding in the second step of qualitative analysis and, subsequently, the allocation to specific heritage value typologies/categories.

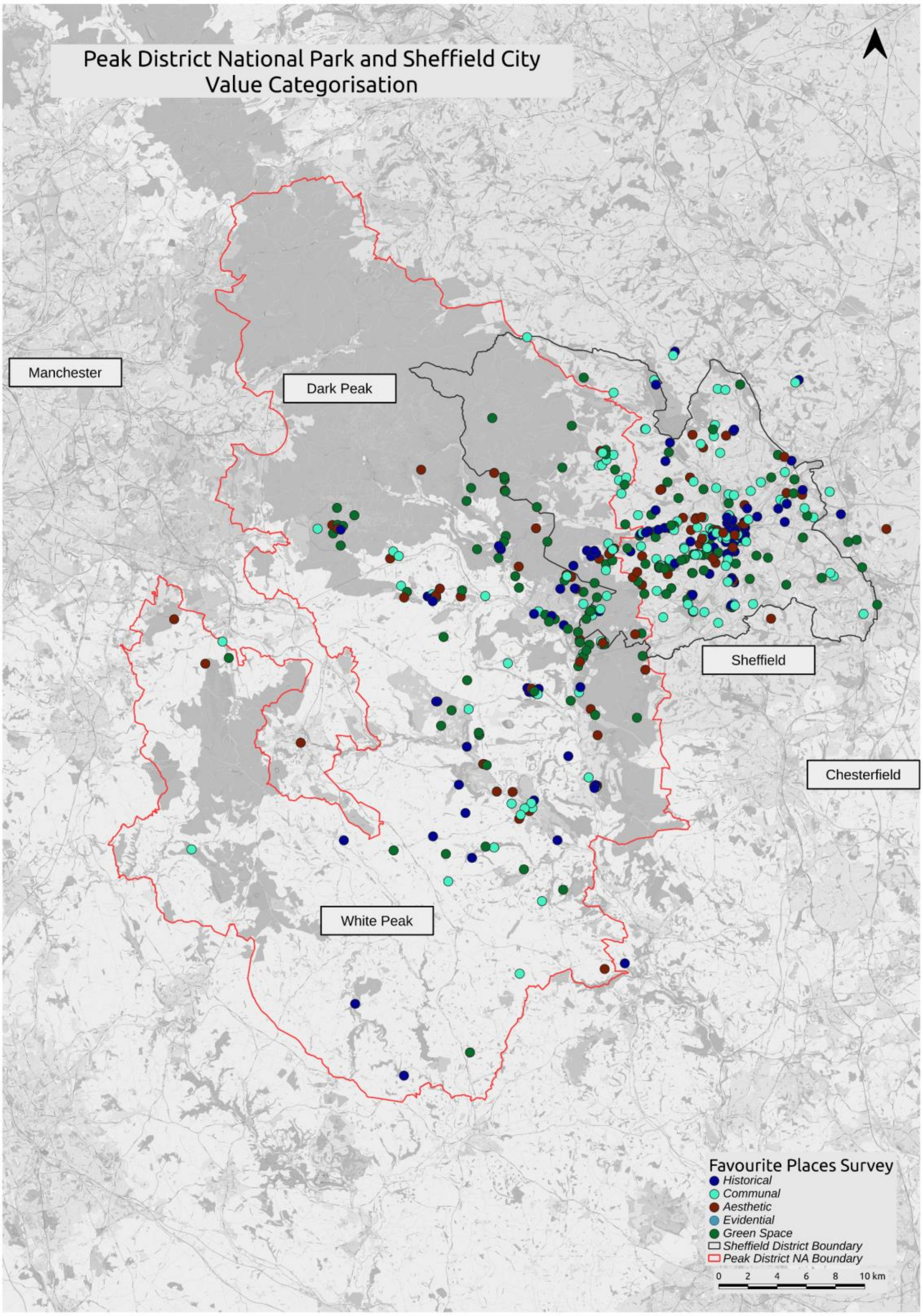

Figure 4-8: Mapped overview of categories across the study areas. Distinct areas show clusters. (Map created in QGIS; data contain OS data © Crown copyright and database right 2022. Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL).

4.4.3 Category visualization

QGIS was used throughout this project. The visualization of favourite places in both study areas can be seen in Figure 4-8. The map is presented as an interactive webmap where the points on the map can be selected and information on the personal connection (stories) provided in the form of a pop-up window. Photographs provided by the participants illustrate the places as an extra layer. Individual values can be selected and presented on the map, for example, only showing locations labelled with an aesthetic value or values for the calming qualities of the landscape. Figure 4-8 provides a background overview of the categories as set out by HE's *Conservation Principles* and the proposed additional categories as a result of this study.

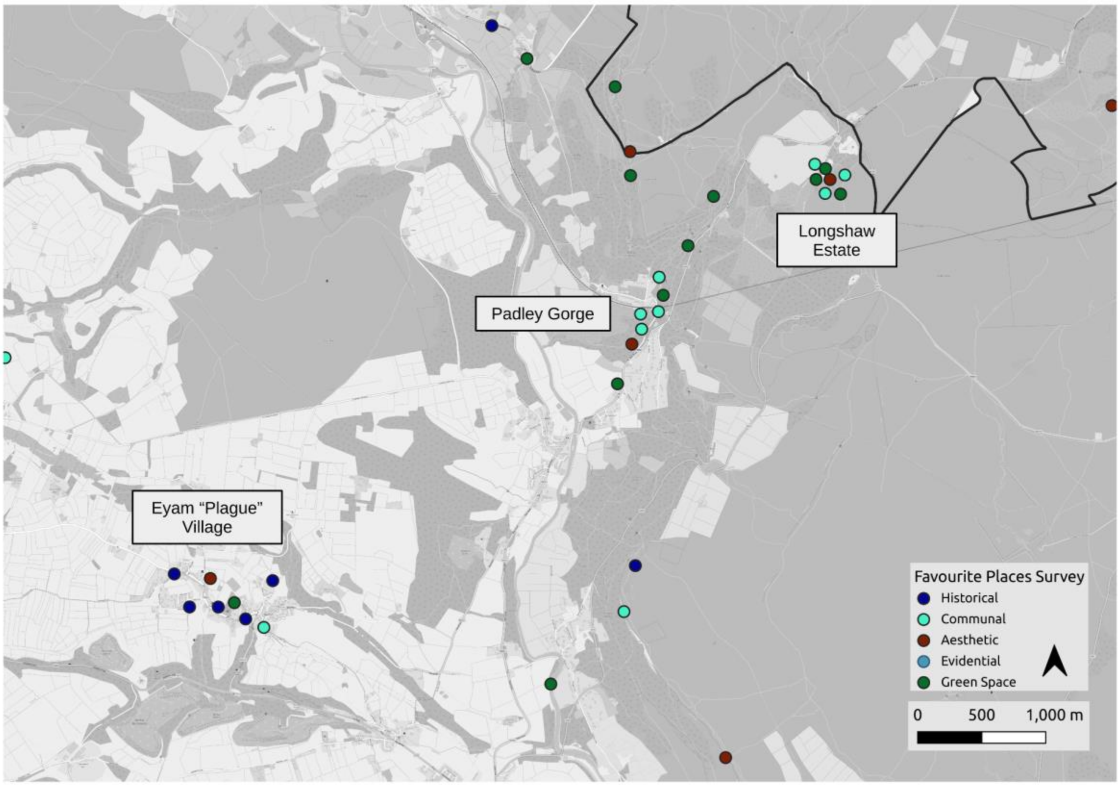

Figure 4-9: Detail of the overview map (Figure 4-8). The plague village Eyam (lower left) is predominantly valued by survey participants for the 'Historical' value (n = 5). Padley Gorge (centre), one of the temperate rainforests of Britain (Shrubsole2022), is valued for the 'Green Space' qualities (n = 4) and 'Communal' values (n = 4), similar to the National Trust's Longshaw Estate (upper right) with recognition given predominantly to 'Green Space' value (n = 3) and 'Communal' value (n = 3). (Map created in QGIS, © OpenStreetMap contributors; data contain OS data © Crown copyright and database right 2022.)

The detailed maps give an indication of the values given to places as identified in this study. The values emerged from the stories based on personal attachment to these places. Figure 4-9 focuses on a central part of the PDNP with the village of Eyam, which is known as the 'plague village'. An outbreak of bubonic plague in 1665 forced the residents of Eyam to isolate themselves under huge hardship to save the surrounding villages from a spread of the disease.[14] 'Historical' value dominates this location (n = 5).

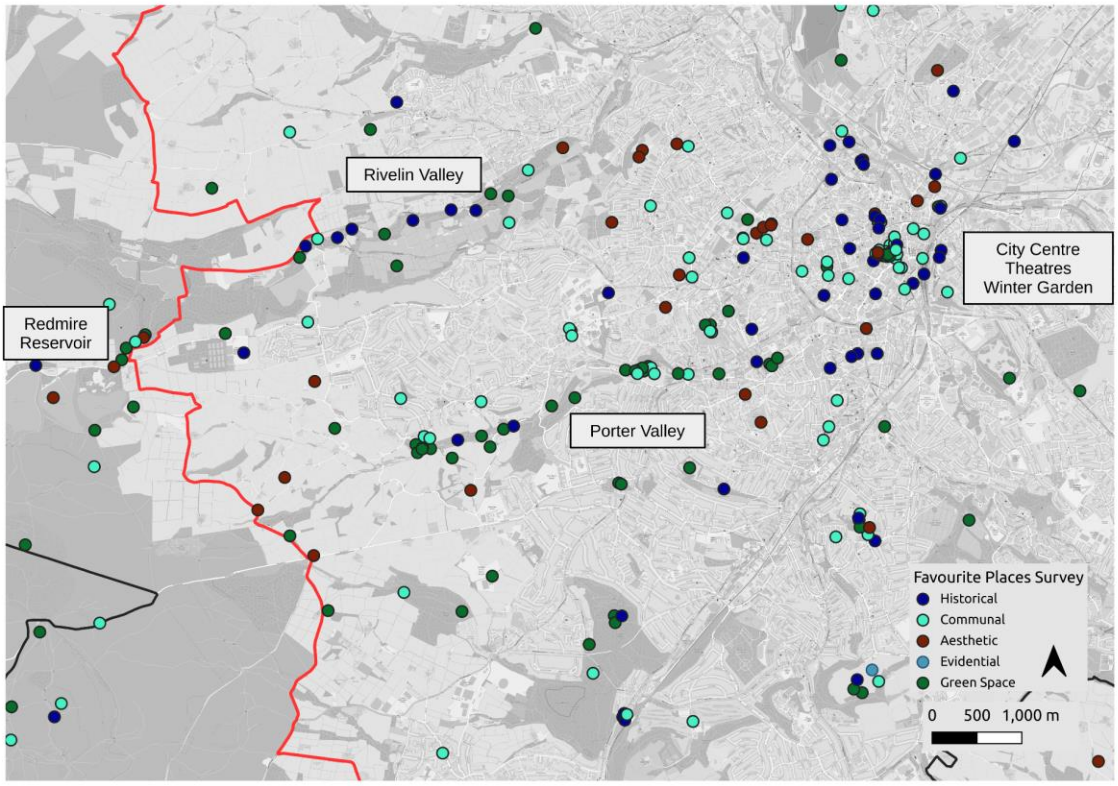

Figure 4-10: Detail of the overview map (Figure 3-8). Notable is the difference between the public perception of the two valleys: Rivelin Valley (upper centre) and the Porter Valley (centre). Rivelin Valley, with its deep early industrial history present in the various sites of 'Wheels' along the river, is predominantly valued for its 'Historical' value (n = 6), while the Porter Valley in the heart of the city is predominantly valued for its 'Green Space' value (n = 14). (Map created in QGIS, © OpenStreetMap contributors; data contain OS data © Crown copyright and database right 2022.)

Notable for the river valleys, leading from the PDNP into Sheffield, is the difference of value-based connection as shown in Figure 4-10. The Rivelin Valley is predominantly favoured for its 'Historical' value which refers to the up to 20 watermills along the watercourse appearing from around 1600[15]. In contrast, the Porter Valley to the south is predominantly favoured for its 'Green Space' value and

[14] https://www.eyamvillage.org.uk/

[15] https://www.joinedupheritagesheffield.org.uk/groups/rivelin-valley-conservation-group

health benefits as a green space in the city centre. A detailed view on the city centre, visualizing categories and subcategories in Figure 4-11 provides insight into the values of residents associated predominantly with 'Arts & Culture' and 'Historical' value (each n = 12).

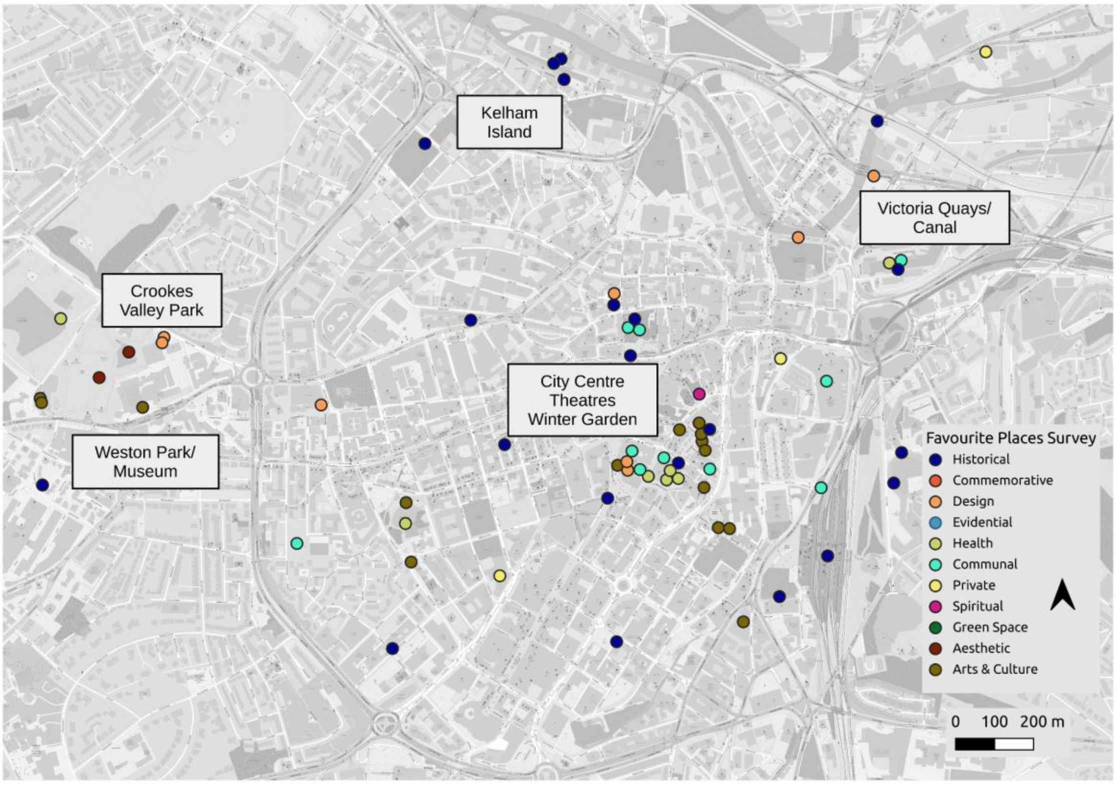

Figure 4-11: Detail of the overview map (Figure 3-8). The legend provides a more fine-grained categorization. 'Historical' (n = 12) and 'Arts & Culture' values (n = 12) dominate. (Map created in QGIS, © OpenStreetMap contributors; data contain OS data © Crown copyright and database right 2022.)

The mapping of values that has emerged from the survey shows that individual, subjective stories have the potential to form an overall value pattern at landscape scale. The value distribution presents the varying and manifold valuation of places, showing how one place can have more than one meaning for the public. Nevertheless, despite a multiplicity of meanings and significances, current techniques allow the assessment, integration and visualization of public perception and place attachment in a form that can provide vital background for planning, alongside (notably and for example) Historic Landscape Characterization or HLC.

4.5 *Limitations, biases and advantages of AI*

It should be pointed out that qualitative research introduces bias at various stages of the research process. Our data collection introduced a limitation and bias as a convenience sample through the means of publication and completion of the survey. We fully acknowledge that marginalized groups have not been specifically explored and would certainly be a study that could follow up, be complementary to and contrast the findings of this paper (see Footnote 4). Both approaches—TM and Manual Observation/Annotation—introduce biases into the process of data analysis. The only automated process that is not influenced by the researcher is the phase where the TM model runs (the computation of the models). TM is based on a model algorithm, which is sometimes treated as a black box because of the complex mathematical ground on which the algorithms and statistical methods are based. However, to make the process as clear and transparent as possible, the choice of model algorithm and the definition of parameters must be documented. Appropriate models can be chosen depending on the analysis and data. Model parameters, such as iteration or number of topics, can be optimized and adapted to control the process.

The final assessment and analysis for the training data[16] is manually observed by the researcher to create a model based on the training data, which can then be applied to new data for an optimization of Topic Modelling in the specific field—in this case, the categorization of heritage values in cultural landscapes. This introduces the human factor, which in turn introduces researcher bias into the methodology, similar to other qualitative research methods (i.e. NVivo or manual coding). Transparency and reproducibility can be achieved by rigorous documentation of the process.

By way of advantage, the use of TM allows topics and themes to be identified that emerge from or are latent within the data without having a preconceived set of codes. This approach eliminates the risk of researcher bias towards the topics

---

[16] Part of the survey dataset that is set apart to train the model for an automatic categorization of new data; test data is the remaining part of the survey dataset that is used to assess the performance of the model.

introduced in the data analysis by preconceived codes. This constitutes a significant change in the way qualitative data can speak for themselves and can reveal patterns within the data, before the topics are then further analysed manually in the process of categorization.

AI offers the opportunity to create a reproducible, repeatable and automated workflow of processes, which would not be possible with various individual human assessors. The routine set-up for the analysis works on the same parameters in every iteration of the process. This makes it more reliable for the categorization of data provided by future surveys to review people-centred values on a rolling basis.

4.6 Discussion

Our methodology builds on approaches tested in previous studies: on research projects focusing on a people–place connection (Dalglish and Leslie 2016; Natural England 2015; Primdahl and Kristensen 2016) and on approaches of other disciplines focusing on computer linguistics that provide methodologies for an application in heritage management (Goerz and Scholz 2010; Sassolini and Cinini 2010; Sporleder 2010). AI has been successfully applied in Heritage Management, but not on a landscape scale (Bordoni, Mele and Sorgente 2016; Condorelli et al. 2020; Fiorucci et al. 2020; Matrone et al. 2020). Our research methodology therefore combines the approach from disciplines such as data science, geography, archaeology and urban planning on a landscape scale. Using Topic Modelling and a Grounded Theory approach to analyse and integrate public survey data as spatial representation into the planning process offers a way to overcome the challenges qualitative data can often represent in terms of practical application.

Themes emerging from stories of place attachment provided by survey participants in the Peak District National Park and the City of Sheffield align well with some of the existing heritage value categorizations set out by Historic England's *Conservation Principles* (Historical, Evidential, Aesthetic, Communal) (65 per cent). Notable, however, is the lack of 'Evidential' value (n = 1) in the perceptions of people with just one case, perhaps due to the lack of documentation on this value subject in a publicly accessible format (e.g., information boards, accessible documentation).

However, ‘Communal’ value, usually not prioritized among other values in the significance assessment process, represents almost one-third of the traditional value categories. Within the City of Sheffield, the valuing of the cultural element of the city centre was notable, for which we propose an additional subcategory of ‘Arts & Culture’ value (which comprised 15 per cent of all ‘Communal’ value). It also became clear that 35 per cent of all personal connections to place were based on experiences during the COVID-19 pandemic and with a strong orientation toward the use of landscape for mental and physical health benefits (Fagerholm et al. 2022). Therefore, we propose the introduction of the category ‘Green Space’ value to bridge the artificial division between natural and cultural heritage. Some 71 per cent of the data in this category showed a close relationship between nature and health.

These results show that single, individual opinions collected in a public online survey can map onto the landscape as heritage values based on people's individual, personal and subjective connections to place. The argument that individual opinions or the public view cannot be considered in the valuation process of landscapes (because it is too difficult, or the data are too diverse to be meaningful) is therefore no longer tenable. We will return to this point in the conclusion. Individual opinions—when aggregated—can create a generalizable pattern of heritage valuation. The detailed view of the two river valleys in Sheffield exemplifies how personal stories can form a coherent pattern of such values. We present this example, therefore, as proof of concept for a method that is suitable for integrating people-centred landscape perception and values within a framework for assessing landscapes within planning and decision-making processes. We also argue for this grounded approach providing the basis for generating entirely new categories or sub-categories of values, rather than trying to force observed values into a predefined framework. While we suggested a bottom-up approach to obtain insight into the perceptions of people's understanding and valuing of the heritage in their everyday environment, the information should be understood as an additional layer of background information in the framework created by Historic England. This dataset will enable us to make decisions shaping people's neighbourhoods on a wider base of information including the crucial dataset of ‘insider knowledge’. While this study provides a snapshot in time, complementary studies can contribute to a completion of the picture and further automation of the process can allow data collection and analysis on a rolling basis for up-to-date datasets (Tenzer 2022; Tenzer and Schofield 2023).

Our research has highlighted potential areas for future work, for example, the automation of the topic clustering and labelling process, including the application of the proposed value categories in a supervised topic modelling approach. This is necessary as the unsupervised model only works with words used in the documents (stories) and not the English Heritage (2008) value categories. The topics are only sometimes sensible and coherent (meaning high coherence and meaningful clustering). In the future, the datasets created from this work could be used to train supervised deep-learning models that can then be used for fully automated labelling and the classification of users' responses.

Furthermore, there is scope to integrate a more refined approach to survey participants and the inclusion of visitors or demographically to explore variance in perception. A more fine-grained study could explore marginalized communities and the integration of heritage from a different perspective. Also, a focus on the visualization of qualitative data could develop a more nuanced representation of value categories in GIS. Finally, developing the theory of perception and value production in different contexts and communities could increase the understanding of what makes landscapes and people's everyday heritage vital for developing a sense of place, identity and belonging.

AI technologies are set to revolutionize the opportunities to understand and present datasets from heritage and archaeology. New tools and methods are currently being developed by a loose community of digital archaeologists, computer linguists and data scientists that, together, have yet to find a more inter- and trans-disciplinary approach to the question of values. However, the trends to cooperate are promising and will bring into the future the methods of various disciplines concerned with the past and with landscape.

4.7 Conclusion

This study has developed a novel method that allows the integration of people's connections with place into the assessment framework of landscape and heritage management. Furthermore, this research has introduced Topic Modelling combined with a Grounded Theory approach with a bias-reduced, time-efficient and repeatable method to interpret and categorize people-centred values of everyday heritage. This innovative approach to qualitative data collection for heritage planning can give local authorities and heritage organizations the opportunity to embed locally held values of heritage within landscape management processes.

The meanings and values people place in the landscapes of their everyday lives are more varied and personal than widely considered in heritage and landscape assessment. Individually held, subjective values form a category for heritage assessment that should not be underestimated for its capability to shape identity, create deep bonds and positively impact place-making [in terms of] more than just anecdotal evidence (Modesto and Waterton 2020). However, this is not about individual opinions; rather it is about gaining a deeper understanding of what drives the development of a sense of place, belonging and identity, and how this can change over time. It is also about recognizing the significance of locally held views and values in creating a more inclusive approach to heritage management than that which exists currently, at least in the Anglophone world. This study shows how diverse meaning and valuing is within communities, but that it can form a distinctive pattern across the landscapes and that this pattern can be both captured and accommodated within the planning process and in heritage management strategies.

Acknowledgement

This project was carried out as a part of an AHRC/UKRI WRoCAH-funded PhD project at the University of York. Grant reference number: AH/R012733/1. The authors want to thank Anna Badcock (PDNPA) and Zac Nellist (SYAS) for their support for this project and help to publish the survey through organizational channels.

Supplementary materials and additional information

Supplementary material for this paper can be found at the end of this thesis in the Appendix section (Appendices for all chapters, **B. Appendices for Chapter 4:**

Supplementary Material 1 – 3, **D. Appendices for Chapter 8: Appendix 2 – 4**, providing the Consent form, Information sheet and JavaScript for Google map embedding). References can be found in the joint reference list at the end of the thesis.

The co-authorship of John Schofield is acknowledged for his contributions to conceptualisation and writing (critical review and editing – 15% share).

Chapter 5:

People and Places:

Towards an Understanding and

Categorisation of Reasons for Place Attachment

– Case Studies from The North of England

Published as:

Tenzer, M., & Schofield, J. (2023). People And Places: Using Topic Modelling to Identify Reasons for Place Attachment – Case Studies from the North of England. *Landscape Research*, 49 (3), 340-358.

DOI: 0.1080/01426397.2023.2289970

Abstract

People develop a sense of place, belonging and identity when a place affords tangible and intangible benefits like security, familiarity, shelter, food, work opportunities, and social interaction. Places form landscapes individually valued by people for these reasons. This paper describes Topic Modelling as a new grounded approach to assessing people's sense of place in a rural landscape affording special qualities for everyday working and living situations – the Peak District National Park, UK. This novel approach is applicable and scalable to any landscape, rural or urban, iconic, or everyday. Results of this study show that significant themes and phenomena not hypothesised at the initial research design stage can emerge from interview data. Examples include pro-environmental behaviours resulting from traditional farming practices, environmental benefits of the drystone-walling tradition, and attitudes towards rewilding initiatives. We argue that such phenomena arise from people's attachment to place and influence their behaviours.

5.1 Introduction

People connect to places for different reasons and can develop a sense of place, belonging and identity through those connections (e.g. Cresswell 2015; Feld and Basso 1996; Jones and Leech 2015; Seamon 2020). This occurs when a place affords tangible and intangible benefits like security and familiarity, shelter, food, work opportunities, social interaction, and a space for well-being. The human geographer Yi-Fu Tuan developed the 'sense of place' concept to theorise connections between people and places (Tuan 1977). Such connections were subsequently developed as the concept of place attachment. Over the past 50 years, place attachment studies have provided theory and methods for assessing the connection between people and places (e.g. Altman and Low 1992; Lewicka 2011). The following will elaborate on (1) place and (2) reasons for attachment.

(1) Place, as a concept, is variously defined across disciplines (e.g. Cresswell 2015; Ingold 1993; Tilley 1994). Places are elements of rich and dynamic historic and contemporary landscape that require various forms of heritage management and planning. Tools, such as Historic Landscape Characterisation (HLC, e.g. Aldred and Fairclough 2003) and Landscape Character Assessment (LCA, e.g. Swanwick 2004), were developed specifically for such planning and management purposes. The focus on integrating people-centred approaches to decision-making for public benefit has only recently entered into management and policymaking but is still dominated by '"objective" outside experts' (Butler 2016, 1). But there are examples indicating a change. For example, in Scotland research conducted by Community Land Scotland with Inherit involved interviewing practitioners and members of the public to explore ways of integrating public opinion meaningfully into the planning and decision-making process (Dalglish 2018, 54-55; Koblet and Purves 2020; Wartmann, Acheson and Purves 2018). Nevertheless, landscapes – and places – are subject to change and development and there is a significant research gap in identifying the impact of change on place attachment and people's perception of the quality of familiar places (Hedblom et al. 2020, 58; see also Hunziker, Buchecker and Hartig 2007). This gap is due mainly to the inherently challenging and time-consuming qualitative data analysis methods that are typically used in such investigations.

(2) We have previously identified the significance of social values held by individual people or communities as reasons for a connection to places and landscapes (Tenzer 2022;

Tenzer and Schofield 2024). Social value is created when the meaning of the qualities of a place are weighed and signified (Williams and Patterson 1999, 142) and can emerge from family or local history, memory, traditions, myths, legends, and beliefs (Jones 2017). Stefaniak, Bilewicz, and Lewicka (2017) showed the development of place identity as the relation of people's personal life stories with the history and past of a place based on ancestry, memories, and traditions, emphasising that people who wish to 'actively engage with place ... come to feel a part of place' (Seamon 2020, 37). In this holistic approach to landscapes, we follow Tilley and Cameron (2017) who recognise the agency of both human and non- human actors and the material world and an emotional, social bond between people and places.

In this paper, we explore what insights can be gained into the meaning and significance of social values through in-depth interviewing of people living and working in the Peak District National Park (PDNP) (UK) with a view to tangible and intangible elements in the landscape that afford strong connections. We categorise these elements and ask how a strong place attachment influences people's behaviours and approaches to the place or landscape? These insights can provide vital background for proactive planning and development, adapting to the needs and visions of people with a strong place attachment to be socially sustainable.

We propose the application of Topic Modelling to interview data as a qualitative assessment methodology, which is time- and resource-efficient and allows the exploration of qualitative data. Artificial Intelligence tools, such as Natural Language Processing and Topic Modelling (TM), have been developed since the late 20th century. The Council of Europe actively encourages the use of these new tools within heritage activities (e.g. Traviglia 2022). However, the deployment of these methods and tools in heritage studies is a recent development and they have yet to be fully integrated into the field of heritage and landscape studies (but see Fiorucci et al. 2020; Matrone et al. 2020; Purves, Koblet and Adams 2022; Bordoni, Mele and Sorgente 2016; Wartmann, Koblet and Purves 2021). We apply TM, which we have previously used to analyse survey data (see Tenzer and Schofield 2024; see also Abram, Mancini and Parker 2020; Cai et al. 2021; Franzosi, Dong and Dong 2022). TM allows the exploration of qualitative unstructured data to reveal themes latent within or emerging from the empirical data without preconceived assumptions. This principle adheres to the Grounded Theory elements underpinning this research (e.g. Charmaz 2006; Odacioglu and Zhang 2022). While the case study focuses on rural areas, this approach is applicable and scalable to any landscape.

5.2 Data source and methodology

5.2.1 Overview of the method

Typically, place attachment is measured using Likert scales, which risks missing the fine nuances shaping attachment (Boley et al. 2021; Brown, Raymond and Corcoran 2015). Social values as a basis for place attachment can be captured in different ways. Surveys have been used elsewhere and provide a broader view on the reasons behind attachment (e.g. Wartmann, Koblet and Purves 2021). For this study, we apply the approach we developed by Tenzer & Schofield (2024) and apply Topic Modelling to interviews with people living and working in the Peak District National Park as a first pre-assessment stage of the data analysis, followed by a second stage of manual observation.

Figure 5-1 demonstrates the steps involved in the combined approach of Topic Modelling and direct observation used in this research. Interviews were collected in person at the places to which interviewees described as having the strongest bonds. Single paragraphs were then treated as text documents and fed into the topic modelling algorithm. This approach allowed the discovery of themes within the data that might otherwise stay undetected or disguised by the researcher’s assumptions and predefined codes.

5.2.2 Study area

The Peak District was designated as the first National Park in the UK in 1951 (Figure 5-2). The PDNP has a history of at least 10,000 years of human occupation and covers an area of 1438 km^2. It has 38,000 residents and receives more than 13 million visitors per year. The PDNPA contains a large number of designated sites, including 2900 listed buildings, 109 conservation areas and 450 scheduled monuments, including prehistoric burial mounds, stone circles, medieval field systems, castles and country houses (PDNPA n.d.). One third of the PDNP are Nature Protected Areas, designated as *Sites of Specific Scientific Interest* (SSSI), Ancient Woodland, *National Nature Reserves* (NNR) and *Local Nature Reserves* (LNR). A specific character feature of the National Park are the drystone walls with a total length of

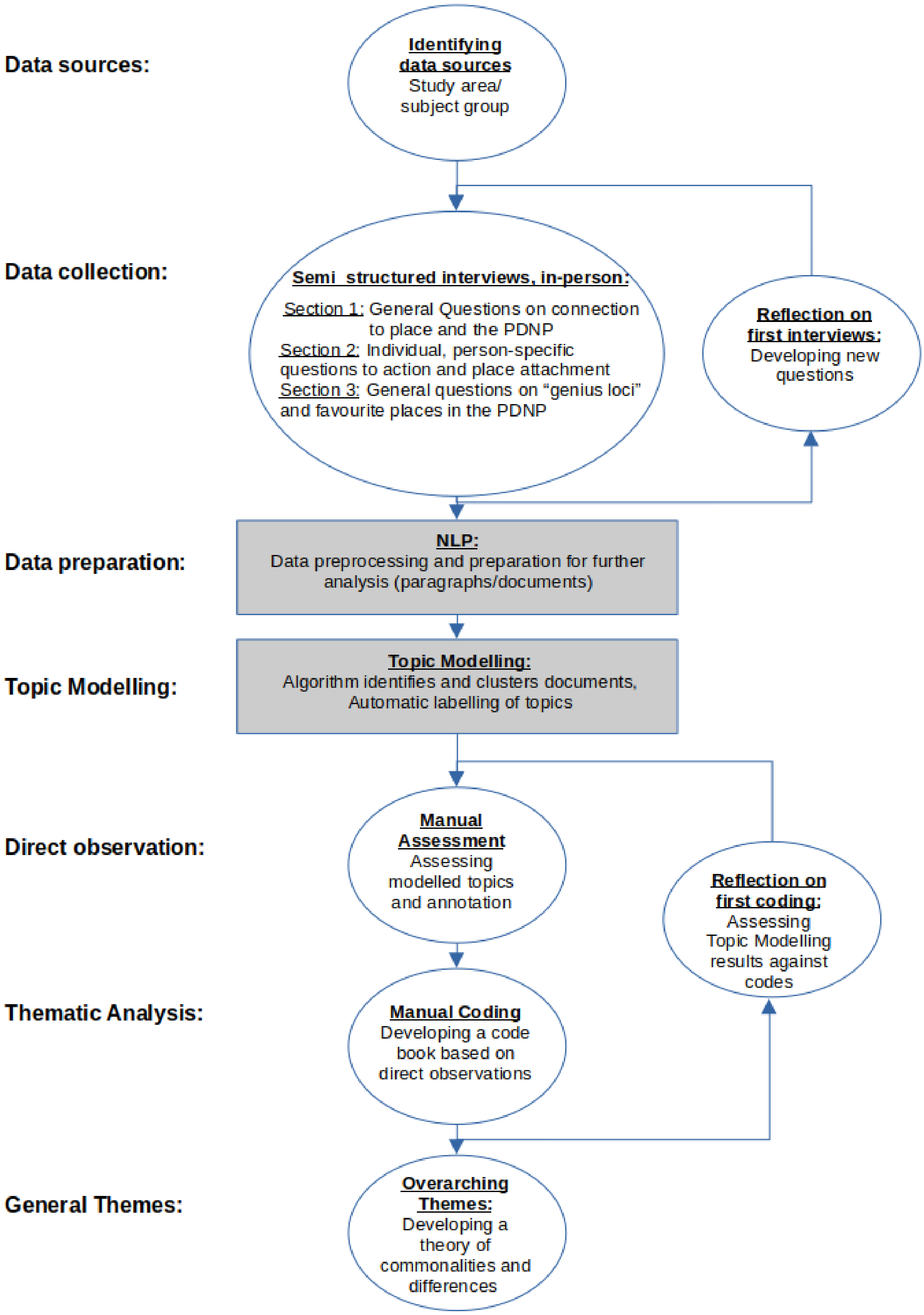

Figure 5-1: A flowchart for the Topic Modelling methodology developed in this research. The aim is to create the topics based on interviews and develop a general observation of landscape perception (Key: white: manual process, grey: automated process).

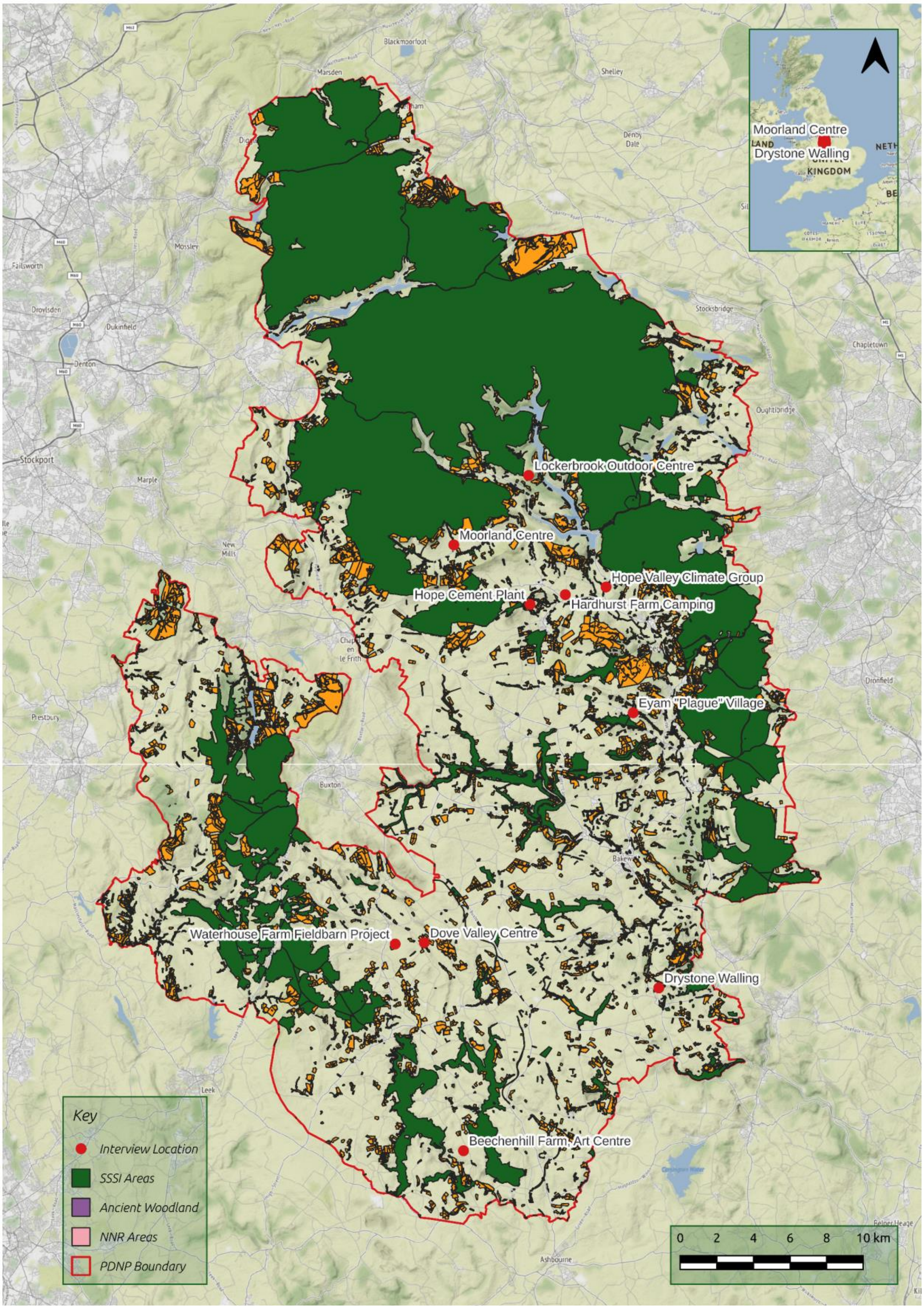

Figure 5-2: Map of the study area: Peak District National Park. The coloured parts of the park show the location of Protected Areas (PA: SSSI, National Nature Reserve, Ancient Woodland). Location of the interview participants are marked in red. Source: PA map data https://designatedsites.naturalengland.org.uk/GreenInfrastructure/Map.aspx. Basemap # Crown copyright and database right 2022, Map tiles by Stamen Design, under CCBY 3.0. Data by OpenStreetMap, under ODbL.

26,000 miles[1].The palimpsest of human impact on the landscape includes traces of industrial action including remnants of the millstone industry and rakes from lead extraction. This quality of open landscape across the PDNP stands in stark contrast to the adjacent industrial cities of Sheffield and Manchester.

5.2.3 Data sources

To assess the landscape factors dominating the perception of people living/working in PDNP, ten in-depth in-person interviews were conducted over 16 months between late 2021 (after COVID-19 restrictions were lifted) and early 2023 (locations in Figure 4-2) (**Supplementary Material 1**[2]). The case study approach was chosen for in-depth data on perception of individuals within a bounded system (Creswell 2017, 96; Flybjerg 2011). This approach completes a triangulation of methodologies developed to extract information on perception (Altman & Low, 1992; Low, 2002), meaning-making and value creation in the study area based on three different data sources: social media data (Tenzer 2022), online survey data (Tenzer & Schofield 2024), and in-depth interview case studies, as presented in this paper. The sampling of participants was based on the Typical Case sampling strategy advised by the PDNPA and from research into place history, providing typical, information-rich examples for the exploration of landscape perception (Creswell 2017; see also Koblet & Purves 2020; Tilley & Cameron 2017; Wit 2013).

The semi-structured interviews were held at the place of residence or work of the interviewees and comprised three parts. Part 1 involved general questions about people's connection to the PDNP and general landscape and local heritage perception. Part 2 of the interviews focused on the specific aspects of living or working in each particular case and were dynamic. Part 3 focused on tangible and intangible elements of the landscape. The interview transcripts ranged in length between 4351 and 12 773 words. The answers were

[1] https://www.peakdistrict.gov.uk/learning-about/news/70-years-of-the-peak-district-national-park/peak-district-facts

[2] Available in **B. Appendices for Chapter 5**.

separated into paragraphs, resulting in a total of 298 separate paragraphs (documents) and forming the dataset for the analysis using TM tools and manual assessment.

5.2.4 Topic Modelling and direct observation

We apply Topic Modelling to interview data as a qualitative assessment methodology, which is time and resource-efficient and allows the exploration of qualitative data as developed and described in Tenzer and Schofield (2024). This algorithm allows an insight into the empirical data that conforms to fundamental elements of Grounded Theory (Charmaz 2006), where data are clustered by emerging themes. This approach allows the discovery of themes within the data that researchers might not have anticipated or discovered (objectifying researcher bias).

For the data analysis, we used the R package textmineR to pre-process the data (including data cleaning and lemmatisation) for TM, following an innovative method developed by Jones (2021; see also Jones, Doane and Attbom 2021). The method implements Latent Dirichlet Allocation (LDA), a statistical method to identify themes based on keywords in the documents (Blei, Ng and Jordan 2003; tqx94 2022). This statistical analysis method clusters text according to themes based on keywords. For a detailed description of the TM methodology, see Tenzer & Schofield 2024). Here we give only a short outline.

LDA clusters the data into a predefined number of topics. The optimum number of topics can only be tested but not determined from the start. Therefore, 60 models were created and assessed using the topic coherence factor (Figure 5-3). The higher this factor, the more association between the words in each topic cluster. Figure 5-4 shows the relationship between number of topics and coherence of terms within topics. We choose *k=37* topics since the curve flattens out beyond, meaning coherence does not increase significantly with a subdivision into more clusters. TM forms themed clusters based on the probabilistic distribution of words over topics and topics over documents, labelled as bi-grams (two closely associated words in a text) (tqx94 2022)[3].These labels are not meaningful as such but give a good indication of the latent themes (**C. Appendices for Chapter 5: Supplementary Material 2**) for manual labelling.

While the application of TM can support the first thematic insight into the data, it cannot and should not replace the direct observation of content and themes (Chang et al. 2009). Therefore, the subsequent manual evaluation and annotation included an assessment and refinement of the topic labels, as described in the next section.

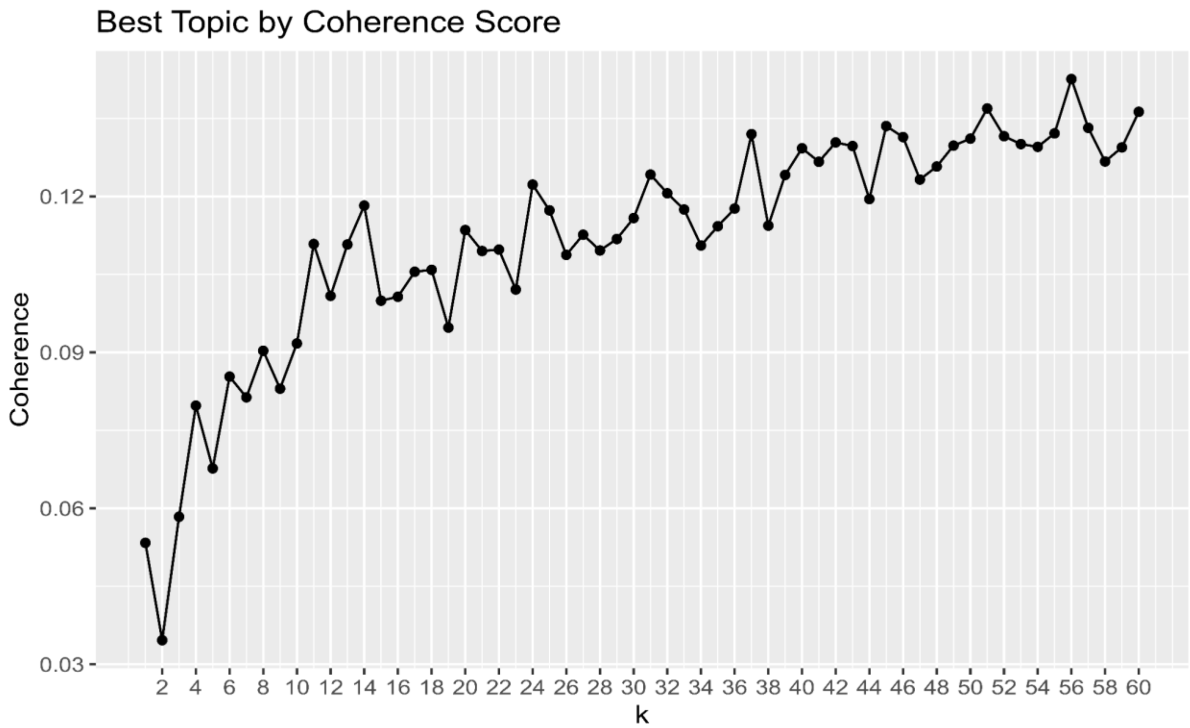

Figure 5-3: Model creation with the Topic Modelling algorithm. Sixty Models were created (k). The best topic coherence in a model can be assessed by the coherence score (association of relevant keywords in the documents and their best fit to each other). High coherence is achieved at 37 topics with a flattening of the curve for more topics.

---

[3] From here, the following terminology will be used: Topics created by the algorithm will be referred to as topic_number and for the label: keyword_keyword. Manually created topics and labels for the code book will be presented in the format: Topic+number and "Code Label". The format: Topic number1/number2 shows the correlation of automatic and manual labels. For example, Topic 16/1 refers to the Topic Model label topic_16 correlated with the manual category 1.

5.3 *Results*

5.3.1 Direct observation and categorisation

A first observation of the model labels gave insights into the themes latent in the data themes (**C. Appendices for Chapter 5: Supplementary Material 2**). For instance: 'stewardship_scheme' and 'tree_plant' pointed towards a connection of management in the PDNP and the subsidies for pro-environmental action; 'plan_permission' and 'farm_building' hinted at a trend towards diversification of farms in the PDNP; 'national_park' and 'people_visit' contained the theme of tourism and challenges with visitors; while 'climate_change' and 'blanket_bog' offered itself for a theme of climate change that also impacts the PDNP. However, the approach of lemmatisation (grouping of inflected word forms, presenting the basic dictionary form) and bi-gram creation (two words closely associated and positioned in the document) also showed problems with the text-based analysis. For example, 'good_dress' (lemma of 'well dressing') pointed to the tradition of well dressing typical for the PDNP. Only with previous knowledge about the questions and answers can such a label be meaningfully interpreted.

As another example, the labels 'hay_meadow' and 'drystone_wall' were translated into the abstract concept of 'Pro-environmental Behaviour – Biodiversity' (Topic 16/1) while 'plan_permission' and 'farm_building' were manually coded into 'Heritage, tangible – connecting to fabric, building material, object' (Topic 32/14) (themes (**C. Appendices for Chapter 5: Supplementary Material 3**). The manual evaluation resulted in 27 sub-codes ranging from personal life history, connection to the past through landscape history, pro-environmental action, and climate change awareness to perception of working conditions in a national park and the advantages and disadvantages of using traditional skills and methods and summarised in seven overarching topics: Pro-environmental Action (PEA) (*n=52*), Challenges (CHA) (*n=25*), Change and Continuation (CAC) (*n=74*), Communities (COM) (*n=13*), People/Place Engagement (PPE) (*n=26*), Place History (PLH) (*n=41*), Landscape Quality (LSQ) (*n=67*) (Figure 5-4). themes **C. Appendices for Chapter 5: Supplementary Material 3** shows how the Topic Modelling and the manual topics correlate and are categorised.

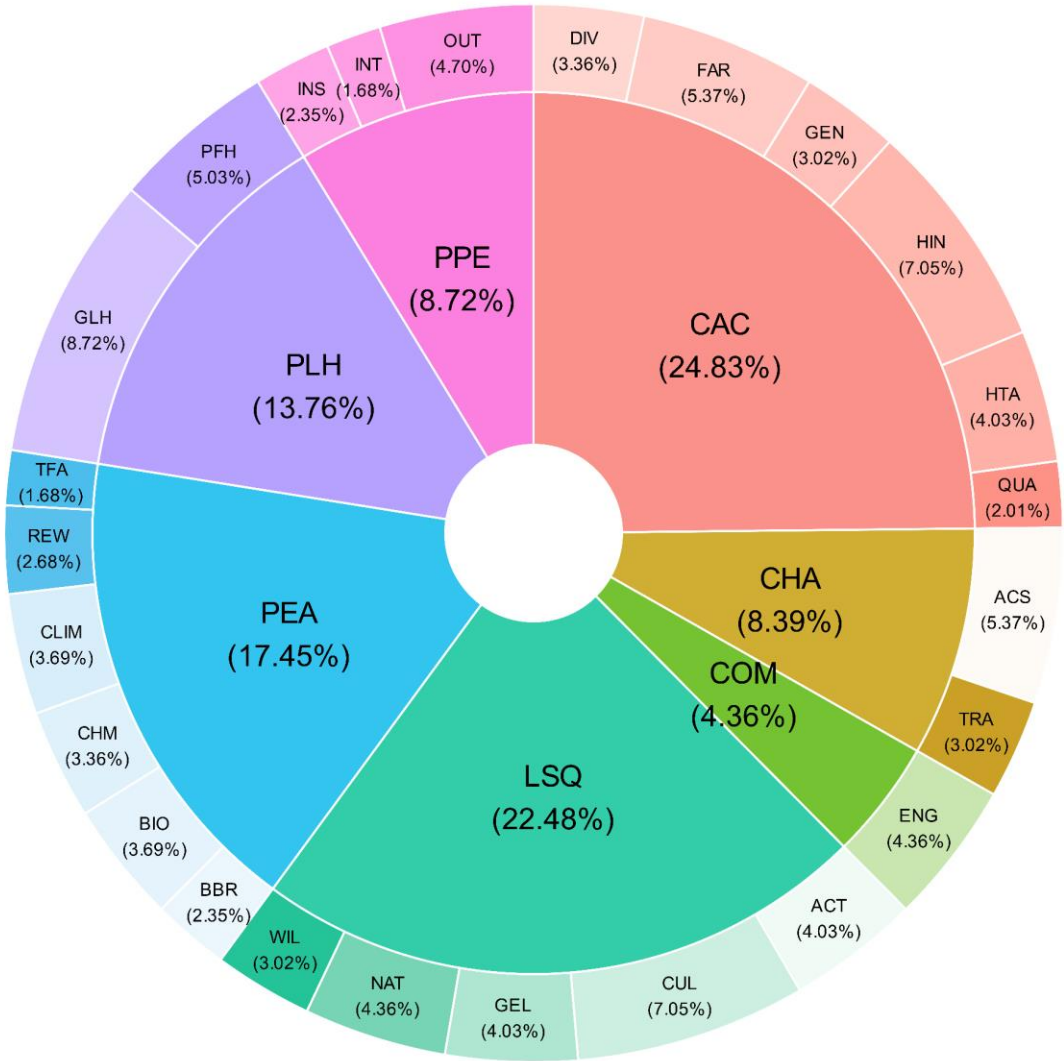

Figure 5-4: Distribution of documents across the topics and subcategories (descriptions provided in Supplementary Material 3. 'Change and Continuation' (CAC) (24.83%) and 'Landscape Quality' (LSQ) (22.48%) dominate the perception of the case study participants. Ot her categories: Pro-environmental Behaviour (PEA), Challenges (CHA), Communities (COM), Place History (PLH), People and Place Engagement (PPE).

5.4 Examples for categories

The following section will provide descriptive examples for the seven overarching categories that were defined based on the Topic Modelling and the subsequent manual categorisation. Figure 5-5 demonstrates the distribution of categories over interviewees.

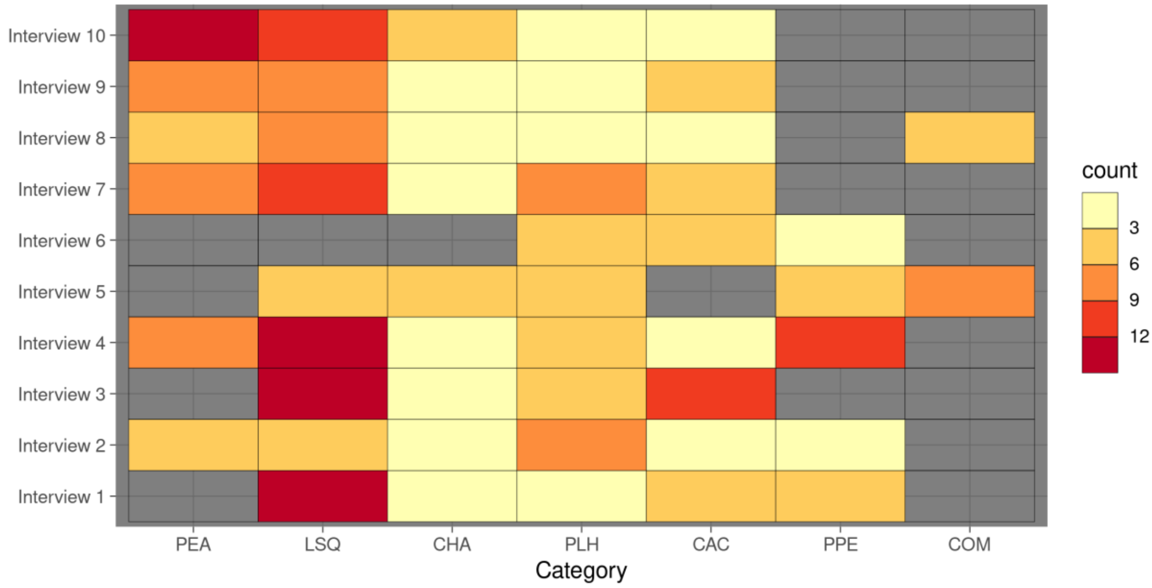

Figure 5-5: Heatmap demonstrating the distribution of categories over the interviews. Each interview contains between three and six categories in varying degrees of intensity. The count shows the number of mentions of each category within each interview.

5.4.1 Pro-environmental action

In Interview 1, the historical skill of drystone walling with a long tradition in the Peak District was the main topic around tangible and intangible heritage, recognising the potential of creating a sense of identity and belonging through an embodied experience of heritage as part of the cultural landscape (Figure 5-6). The extraordinary qualities and benefits of this form of heritage were mentioned in the first interview with the longest-serving drystone waller, Trevor. He mentioned that the walls provide shelter for many forms of wildlife from insects and spiders to small mammals. However, other interviews (particularly Interview 10) highlighted that drystone walls are also beneficial in confining and limiting the damage from wildfires. Additionally, drystone walls are artificial barriers for surface water flood prevention. Similarly, the reinstatement of the blanket bog in the north of the Peak District contributes to the water quality and so to the eco-system services of the region (Interview 9 and 10) (Figure 5-6).

Similar pro-environmental benefits of historical features emerged during Interview 2, mentioning the re-introduction of such walls and hedges for the subdivision of fields into smaller units. Such boundaries were removed over the past century but are still visible on historical maps. Reinstating boundaries benefits wildlife, as mentioned above. Also, the re-introduction of biodiverse plant species, for instance, traditional apple species, orchards, and wildflower meadows, emphasised the need to promote biodiversity for a sustainable future. Initiatives led by engaged farmers support the exchange of wildflower seeds to extend the gene pool of plant species (Interview 2), and the reason for diminishing landscapes lies in the traditional way of farming:

> I think the other thing is there were lots of wildflower meadows which were farmed environment, a managed environment. And I think we have lost over 90% of them. And most of that again was in pursuit of cash because farmers were encouraged to improve fields and improving means, getting rid of all the wildflowers. (Interview 9)

Figure 5-6: Typical character of the Peak District National Park landscape with the small village of Castleton in the background, Peveril Castle to the right and the typical drystone walls and field barns. Source: photo by M. Tenzer.

A closely related form of pro-environmental action – renaturalisation of former pastureland (Interview 7) – showed a reorientation of PDNP farming community members.

The topic of renaturalisation, rewilding and reforestation, as sub-codes of this category, highlighted the issues around using the term ‘rewilding’, which was unanimously found to be opaque and misleading. All interview participants saw the need to introduce more biodiverse and natural landscapes, but the term ‘rewilding’ triggered an immediate negative attitude towards the concept:

> Not too keen on some of the ideas. We managed without wolves and things ... I don’t like the idea of like wild boars running through forest particularly. Is it really necessary? (Interview 6)

> The difficulty is that rewilding to a farmer or a landowner sounds very much like abandonment. And agricultural abandonment has been seen as really, really bad in the farming community. (Interview 8)

Figure 5-7: Black Hill triangulation point in the blanket bog area, Dark Peak moorland, part of the Pennine Way long distance trail and part of the ‘Moors for the Future’ partnership for the restoration of blanket bog. Source: photo by M. Tenzer.

A question raised by an interviewee asked which point of time to choose for the reinstatement of the 'natural' state:

> I am involved in rewilding activities and there is a big issue about: so, what is the natural state of the landscape? And I don't think there is an answer to that... Do we want it like it was 200 years ago, 2,000 years ago or 20,000 years ago? So, it is not static. And whatever we do, we are going to interfere with nature. (Interview 8)

5.4.2 Change and Continuation

These Pro-environmental Action topics overlap with the theme of Change and Continuation and the sub-code of traditional farming methods. On the one hand, these methods can create an issue with their opposition to innovative thinking, with practices negatively impacting the landscape. On the other hand, the re-introduction of hedges provides an example of an environmentally positive traditional farming feature, contributing to biodiversity. Small acreage (Interview 6) and pressures from market prices for farming produce led to a continuation of improving the pasture by spraying (Interview 6 and 7) and pressures of overgrazing (Interview 7):

> No, we are obviously interested in doing things like that [alternative farming methods] but at the moment we haven't done anything like that because we only have a small acreage most of the land is in use constantly. It is like pastureland, so the sheep graze it all the time. If you know what I mean, we don't have any that we could set aside for wild meadows or anything like that. (Interview 6)

However, changing farming methods can also contribute to a changing landscape. For instance, increasingly, farms are owned by people who are not farming the land. Such land tends to be rented out to farmers, extending businesses on an industrial scale with consequences for work processes. To be more efficient, such farms introduce, for example, larger machinery, demanding a widening of gates and access ways (Interview 3). However, changing conditions can also impact positively, as in the example of the Hope Cement Plant. Vast areas of seemingly devastating quarrying activities to provide minerals for cement production are used as an opportunity to reinstate the exploited areas as renaturalised areas with rich biodiversity (Figure 5-8 and Figure 5-9).

Outside the farming and industrial context, Change and Continuation were dominated by the strong influence of intangible and tangible heritage as reasons why people connect to landscapes. For example, keeping traditions alive, such as well dressing, the Castleton Garland, and sheepdog trials proves important for the coherence of communities and the attraction of visitors. Also, traditional skills were important to the

Figure 5-8: Cement plant quarry adjacent to the north of the plant, current state of work area. Source: photo by M. Tenzer.

interviewees. Interviewees actively work against the loss of skills and traditions, for example, with the creation of a culture centre (Interviews 2 and 7) and the teaching of drystone walling (Interview 1):

We have to preserve some of those...you know...those old traditions, we must. Or we have lost them, and it is nice to see some of the older folk, like myself, teaching the young ones that are interested. (Interview 1)

Figure 5-9: Cement Plant building and the former area of mineral extraction now reinstated with a vegetation of mature small trees and shrubs. Example for the renaturalisation of the quarry area as envisaged for the remainder of the extraction areas. Source: photo by M. Tenzer.

5.4.3 Place History

In view of the historical dimension influencing place attachment coded as Place History, we identified a strong trend in all interviews – independent of whether the interviewee lived or worked in the Peak District or if the interviewee was born or moved into the area – to develop either a strong connection to place through family history or a strong interest in the local history. This was found in the lifelong connection to the landscape or particular characteristic features, such as drystone walls (Interview 1) (Figure 5-6). As another example, there was a deep connection to the 'plague village' of Eyam for one interviewee whose family history can be traced back to one of the survivors of the plague – creating deep roots to the location (Interview 5):

> ... it is the landscape, we have to keep it as it was and try to preserve a lot... we must respect our elders, what they did... But, I still say, it was better then as it is now. (Interview 1)

Contrasting attitudes become apparent when we compare this with residents who moved in, describing the landscape as 'frozen' (Interview 9) or preserved 'in aspic' (Interview 7), referring to the aspect of 'dynamic' landscapes (Interview 8). There is a notable need to accept changing and dynamic landscapes. However, changes and developments in infrastructure and the built environment are identified as having a negative effect on the quality of landscapes in the National Park:

> You know the motorway that crosses the South Pennines just north of the national park boundary. You wouldn't look at that and say: Well that has been a marvellous addition to the landscape, wasn't it? But you would look at Howden Reservoir and most people visit the Upper Derwent valley because it has reservoirs in it and they find this an attractive feature... if we hopped on a hundred years, 200 years we wouldn't look back at the M62 and say: "What a fine piece of industrial architecture that is. (Interview 10)

Comparably, moving into the National Park from outside led to a strong interest in local history, especially the specific history of people's houses of residence and the immediate area around them. Residents developed a keen interest, deep knowledge, and active engagement in the local history – a deep sense of belonging (Interviews 2 and 7).

Correlating the category Change and Continuation with Place History trends showed that participants with a long family connection to the Peak District were more in favour of continuation and preservation of the status quo (Interviews 1, 5 and 6), while interviewees with a shorter period of residence or people working in the National Park suggested that they were more open to change (Interviews 7, 8, 9 and 10). Such change included the diversification of farms, including holiday cottages, culture centres and renaturalisation action (Interviews 2, 7 and 8), a strong awareness of climate change and loss of biodiversity (Interviews 4, 9 and 10), and the associated need for pro-environmental action.

5.4.4 *Landscape Quality*

The strong connection to place was also visible in the perception of Landscape Quality with regard to the familiar, aesthetic and natural qualities. The landscape character proved to have potential for building a strong sense of 'home':

> But coming back into Ashbourne, into the south. As soon as I saw these hills...ah...I felt at home. It was the landscape. (Interview 1)

Opposed to the former concept is the perception of 'man-made' (sic.) versus 'wilderness'. All interviewees acknowledged the notion that the whole of the National Park landscape has in one or another form been impacted by past human action, despite the seeming 'wilderness' of some parts of the landscape:

> So, the major contribution that it [the national park] has made is that we have the last bits of wilderness England. From a people perspective that is actually quite good for our mental health. Because you can't get anywhere in this country where people say: look, this is a wilderness! ... It certainly looks like a wilderness in January in the snow. (Interview 9)

This, again, overlaps with the strong connection to the history of the National Park and the aspect of a living and working landscape:

> So, our national parks ... tend to be looked upon as being up in some conservation of wilderness. And the Peak District and the South Pennines are very much a varied, managed landscape. And I sort of quite like that, and I think a lot of people who visit this area like that. They like that industrial heritage aspect that surrounds it. That story-telling of past human activity which is not quite as strong in other landscapes in England. (Interview 10)

> I think knowing that we have to feed ... there are so many demands on the Peak District. It is not a park in the sense of a wilderness. It is a lived ... a living landscape. Living and working landscape. But it could be ... it continues to be for me the lungs of Sheffield and we need to protect it ... But I would also like to see more bold decisions around land use (Interview 4)

5.4.5 Communities

A strong reason for place attachment could be found in the close-knit community as shown in the example of Eyam:

> Or I have just grown tomatoes for instance, ... I had about twelve or fourteen left over ... and I would put them at the gate in a box with a note on "free to good home" and they all went within two days. And so we share things across community as well. (Interview 5)

5.4.6 Challenges

The interviews gave insight into the issues of residents in the National Park, mostly associated with loss of local housing to holiday and second homes and an associated loss of community (Interviews 3, 5, 6 and 8) and pressures from the increased influx of visitors (Interviews 3, 5, 6, 9 and 10) with damage to local landscapes, disturbance of wildlife and farmstock and daily life of residents.

5.4.7 People and Place Engagement

This category showed the connection between activities that connect people and places. One aspect mentioned in this regard was the stories associated with places that reinforced the connection:

> ...so, I think if we would use the landscape to tell the story I think that then gives us ... it's a story of place that is ... I think that is important. (Interview 4)

In summary, the interview excerpts show the sometimes opposing attitudes amongst people living and working in the Peak District and the different values that people afford to places within its landscape (Figure 5-6). Manifold demands and the diverse needs of people are putting pressure on the National Park, but they also offer opportunities to see the places through the eyes of the residents, to improve the quality of place and lives. After reflecting on the limitations and biases, we will discuss these results and their meaning.

5.5 Discussion

5.5.1 General discussion of results

The connection between people and places has been subject of research with a practical focus on landscape characterisation(e.g. Dalglish and Leslie 2016; Koblet and Purves 2020; Primdahl and Kristensen 2016), and geographically on national parks (e.g.Maguire 2017; Petrova, Čihař and Bouzarovski 2011; Ramkissoon, Weiler and Smith 2012), using interviews as a data source (e.g. Polfliet 2020; Wartmann, Acheson and Purves 2018; de Wit 2013) and deploying AI tools for qualitative data analysis (e.g. Goerz and Scholz 2010; Purves, Koblet and Adams 2022; Sassolini and Cinini 2010; Sporleder 2010). However, approaches in heritage studies have so far not focused on landscape assessments (e.g. Fiorucci et al. 2020; Matrone et al. 2020; Bordoni, Mele and Sorgente 2016). Our research builds on the previous work and combines elements of previous approaches. This paper presents one part of a wider research project exploring social value in the views on place attachment using social media (Tenzer 2022), surveys (Tenzer and Schofield 2024) and in-depth interviews to get both broad and deep insights into factors for place attachment on a landscape scale that can provide background for socially sustainable, inclusive, and transparent management of cultural landscapes.

For this study on TM, we used ten in-depth interviews with people living and working in the PDNP. Case studies provide a deep insight into occurrences that can only be achieved by analysing the particular, 'to draw otherwise inaccessible conclusions' on the perception of specific phenomena (in Creswell 2017, 99; Flybjerg 2011).

The Council of Europe advises testing the deployment of AI in all sectors (Traviglia 2022), and we utilised TM (after Jones 2021; Jones, Doane and Attbom 2021) as a specific application. This tool enabled us to extract themes latent within the empirical data to set a framework or structure using labels based on the words used in the interviews for the subsequent manual assessment and categorisation of the data. With this approach, we adhere to Grounded Theory elements (Charmaz 2006). This supports the discovery of themes emerging from the data that might otherwise not be apparent.

From the initial TM clusters and labels, we inferred seven overarching themes with 27 subcategories of reasons for the connection between the interviewees and places. The analysis showed that ‘Change and Continuation’ (24.83%) and ‘Landscape Quality’ (22.48%) dominated the perception of people living and working in the PDNP. The strong connection between the people and the places was evident from the close-knit community of villagers in Eyam and the community centre for local history, landscape and knowledge education. The interviewees were strongly attached to their places in the PDNP through their love for the landscape, which is partly created through the aesthetic qualities and the seeming ‘wilderness’ and, at the same time, the human traces and achievements inscribed in the historic landscape. Main insights from the data showed that lifelong residents are more resistant to change than long-term residents who moved in (e.g. changing farming methods, change in local population, renaturalisation). All interviewees were aware of pressing problems such as climate change and biodiversity loss and demonstrated action for nature conservation (benefits of drystone walls for wildlife, planting old apple species and hedges, preserving blanket bog). Despite these positive attitudes towards the various places in the PDNP, challenges are part of the life and work of the interviewees, as shown by issues such as increasing footfall of visitors and associated damage to the environment, loss of community coherence through second-home ownership, and negative impact of traditional farming practices. While we have demonstrated that a larger number of individually held social values create patterns across landscapes based on the same values existing across areas or through the application of different values to one place (Tenzer and Schofield 2024), this paper deepens the insight into the manifold values of individual people.

These deep insights support the understanding of reasons for people’s connection to the places where they live or work in the PDNP, which correlate with the findings of the broader approach through survey data as detailed in Tenzer and Schofield (2024). These insights from various data sources can provide essential insider knowledge and background for planning and decision-making in the PDNP, but also beyond this landscape, in other environments, for example, in urban areas or conservation areas. The sample size used in this paper can be scaled up based on time- and resource-efficient TM clustering.

AI tools and methods have applications in a wide range of disciplines. Adoption and adaptation into the field of heritage studies is in its early stages of development. Inter- and transdisciplinary collaboration can aid professionals and practitioners in developing tools for the integration of social values and engagement of the public for a more transparent, inclusive, and socially sustainable planning and management of living and working landscapes.

5.6 Limitations, bias, and the advantage of Topic Modelling

Bias is introduced in various stages of the project, e.g. in the selection of interview partners, the questions, the manual assessment process and the choice of algorithm during the analysis. We fully acknowledge that this selection of interview partners and questions[4] limits the wider representational qualities of this study, but as this is part of a wider project exploring place attachment and social values in the study area, the case studies were intended to explore a few typical cases in-depth as opposed to the study where a larger survey provided a shallower but wider overview of the reasons for place attachment (Tenzer and Schofield 2024).

TM was chosen as a method to explore the data in a first assumption-free approach. The theme-based clustering of LDA is based on statistical principles. In other words, the researcher does not influence the outcome of the computational phase and the method is, therefore, more objective and replicable than manual labelling using predefined codes. As this step is based on complex mathematics and statistics, it is often treated as a black box. Therefore, we need to document the choice of model algorithm and parameter settings, such as number of topics and iterations.

The manual observation and categorisation of the data in the second stage of the process introduced further bias in this analysis. While this process was based on the assumption- free approach through TM, the human factor introduces subjectivity and researcher bias at this stage. Documentation of this process can ensure rigorous and transparent procedures during this step.

[4] Semi-structured interviews used the same “general questions” (part 1) and “other questions” (part 3). Part 2 of the interviews focused on the specific cases and associated themes. This free conversational part introduced researcher bias and themes linking to other interviews, e.g., Did they use traditional farming methods? What did they think about quarries, drystone walls, wildflower meadows, rewilding/renaturalisation?

While the interview dataset is too small for reliable statistical analysis, TM has advantages that can bring new insights and perspectives that would otherwise perhaps not be discovered. It allows empirical data to 'speak for themselves' (Tenzer and Schofield 2024). Furthermore, it enables researchers to scale their research by applying an algorithm and workflow tested on a small dataset to larger datasets (see also Tenzer and Schofield 2024). Such analysis can subsequently be used on a rolling basis reviewing similar data for future analysis. This process can help under-stand changing attitudes towards places over time as a longitudinal study.

5.7 Conclusion

This paper began with the recognition that case studies provide deep insight, adding 'insider' knowledge as background for decision-making and planning, which is still dominated by 'objective' outsider assessments. Sustainable development depends on buy-in of residents and communities through recognising people's visions, demands and needs. People-centred approaches offer opportunities for conversation and collaboration to increase the qualities of place and strengthen the appreciation of places.

To gain insight into themes latent in the interviews – following principles of *Grounded Theory* – we applied Topic Modelling to cluster and label topics emerging from the empirical data. This approach provided a framework from which we developed categories of social values to explore and analyse reasons for developing a sense of place, identity and belonging. The discussion around socially sustainable change and development can be aided by active reflection on the historic landscape and help to develop a better and more nuanced understanding of social values related to working and living landscapes in the agenda-setting and planning of the historic environment. Understanding the reasons for people's practical and emotional connection to places can foster care for the natural and cultural environment, promote inclusivity, and enable socially sustainable landscape management that does not alienate people but recognises and understands place attachment.

Acknowledgement

The authors want to thank Anna Badcock and Catherine Parker Heath (Peak District National

Park Authority, UK) for their support for this project and identifying interview participants in the National Park.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This project was carried out as a part of an AHRC/UKRI WRoCAH-funded PhD project at the University of York. Grant reference [number: AH/R012733/1].

Additional information

The Appendix provides further information for the ethical and GDPR requirements for this research dealing with personal data. Appendices for this chapter can be found in **C. Appendices for Chapter 5: Supplementary Material 2** and **D. Appendices for Chapter 8: Appendices 3 and 4** with Consent form and Information sheet for participants.

The co-authorship of John Schofield is acknowledged for his contributions to writing (critical review and editing – 10% share).

Chapter 6:

Social Landscape Characterisation: a people-centred, place-based approach to inclusive and transparent heritage and landscape management

Published as:

Tenzer, M. (2023). Social landscape Characterisation: a people-centred, place-based approach to inclusive and transparent heritage and landscape management. *International Journal of Heritage Studies 30 (3)*, pp. 269-284.

[DOI: 10.1080/13527258.2023.2289424](https://doi.org/10.1080/13527258.2023.2289424)

Abstract

Landscapes are composed of physical places, affording meaning-making and value creation from everyday heritage based on personal experiences, life histories, memories, traditions and heritage practices. Individually held values form the basis for attachment and connection between people and places. Place attachment develops into a sense of place, belonging and identity. Despite the Burra Charter and Faro Convention's aspiration to include people in the assessment process, individual, subjective or emotional connections to place are often overlooked within heritage decision-making. When places are altered, neglected or damaged, such connections can be lost, and the quality of place diminished. Most changes to landscapes happen as part of the planning process, which is not currently able to account for individual connections but based on views expressed in the language of the Authorised Heritage Discourse (AHD). This paper presents a method to meaningfully integrate insider or individual knowledge into the framework of local planning and decision-making while at the same time addressing subtleties and fluidity of such personal views. The people and place-centred method of Social Landscape Characterisation collects, analyses and visualises invisible or hidden value communities based on the same meaning (category value) or location (place value) as shared values across wider landscapes.

6.1 Introduction: everyday heritage and living landscapes

Everyday heritage consists of a material world – landscapes, buildings, places and objects – that provides the setting for activities and experiences in daily life. People perceive this outer world as an individual or communal experience. But only when this information is processed in the inner world of ideas and cultural imprints do such mundane tangible objects become imbued with meaning (Jacobs 2006). Meaning is formed based on beliefs, traditions, legends, myths, events and local and life histories. When places are imbued with meaning, people begin to value these places for the qualities they afford (Williams and Patterson 1999, 142). Some are consciously experienced and shared by a community, while others are more of an affectionate connection like a feeling difficult to grasp and even more difficult to express. For example, Historic England recognises that 'social value of places are not always recognised by those who share them' (English Heritage 2008, 32). The strong bond people have towards places can be measured through brain activities (Gatersleben et al. 2020).

Internationally and on different levels of heritage protection including designations, shared values as 'social values' or 'communal values' form part of official assessment strategies (English Heritage 2008, 32; see also Johnston 2017; 2023) – from local planning to World Heritage Sites – for everyday landscapes and the elements that constitute them. Social values are currently only defined through a group or community – as shared, negotiated and agreed upon in discursive methods – and these values are slowly being accepted and integrated into the heritage assessment process since their introduction with the Burra Charter (ICOMOS 2013) and the Faro Convention (Convention on the Value of Heritage for the Society (Council of Europe 2005)). However, individually held values are often not recognised and therefore lose out to the expert assessment which is supported by robust data and defined as the AHD (Smith 2006; see also Avrami and Mason 2019; Jones 2017). Social values are typically undervalued in official heritage assessments compared to other values and seen as less authentic because they are 'less capable of constructing a logically consistent and convincing narrative' (Wagenaar, Rodenberg and Rutgers 2023, 1). It becomes even more challenging when an attempt is made to create such a consistent narrative from individual values or the personal connection of people to place and heritage. Social values across the everyday landscape remain hidden and unknown, despite having the potential of forming value patterns that are shared by non-related members of value communities, as will be demonstrated in the case study (below). Such communities are related through their connection and attachment to the same place – most likely without ever meeting or knowing each other. Identifying value patterns as a shared and consistent narrative can be achieved when the place, location or subject/value category come more sharply into focus. Ethnographic methods of enquiry have been shown to provide a deep insight into relationships to specific locations based on a synthesis of individually held values or opinions on a subject (Low 2002; Maguire 2017).

Following Piaget's concept of individual constructivism (Piaget 2013) and Jacobs (2006, 9) development of landscape perception, 'mindscape' form as 'individual values, judgements, feelings and meanings'. 'Powerscapes' on the other hand are influenced by culture and rules and are socially negotiated involving accepted behaviour and codes; while 'matterscape' is the 'material reality, described as a system of facts on which laws of nature apply'. According to Jacobs (2006, 187) landscapes have triggers of natural qualities that allow the emotional response and creation of meaning and connection to people with a predisposition for such qualities. However, when such a valuation of place is then integrated into the current system of heritage assessment, this subjective and individual process can be limited and distorted by the need for identifying a community with shared values, limited by the discourse and biased by the dominant voices in the group or the assessment object. This raises the question: how can local authorities identify significance of everyday heritage and what matters most in people's daily lives and environment on an individual basis? And how can this insider or individual knowledge be meaningfully integrated into the framework of local planning and decision-making?

This paper provides an overview of the current approaches to value assessment in the historic environment in an international context and presents challenges of a people-centred approach to the inclusion of such values into the planning and decision-making process. I will then give a UK- based example of new research that attempts to find ways to include people's views more easily into the process. The developed methodology can be applied internationally and scaled up or down according to need. While Dalglish and Leslie (2016, 217) advocated for a meaningful integration of individually held values into planning considerations, which was difficult to achieve with conventional techniques, Artificial

Intelligence (AI) tools offer new applications and levels of capability. I present a project that has collected, analysed and visualised individually held values in a digital form using AI tools, such as Natural Language Processing and Topic Modelling, and GIS software. In the presented case study, I propose the integration of the resulting Social Landscape Characterisation maps, based on individually held values and subsequently categorised as shared value groups based on location or value category, into the framework of existing heritage and landscape management datasets.

6.2 Heritage assessment strategies and value communities

People's connection to places is complex, difficult to articulate and the format of expressions often not appropriate for inclusion into data sets held at local authorities to facilitate change and development (Common Ground 1996, 2006; Johnston 2023). Historic England's Conservation Principles (English Heritage 2008) identifies four value categories: historical, evidential, aesthetic and communal (see also Emerick 2014, 2016). This approach has been analysed and extended as part of this research elsewhere (Tenzer 2022; Tenzer and Schofield 2023, 2024). As a subcategory of communal value, the Conservation Principles recognise that 'social values of places … may only be articulated when the future of a place is threatened' (English Heritage 2008, 32). In such a case, the potential loss of a heritage asset can trigger the formation of an interest group. Other groups or communities are identified through a shared tradition (see Johnston 2023). Nonetheless, the definition and final assessment of significance of tangible or intangible heritage nearly always rests with the expert.

Defining significance of the historic environment focuses on the assessment of professionals and experts expressing values of place in a language adhering to the AHD. This approach is based on communal values as shared values of the public. For example, in Australia, Johnston notes that the AHC (Australian Heritage Council 2009, 6) – aligning with Historic England's communal values (English Heritage 2008) – frames social value as:

> 'The necessity for the social value to be a shared value […] arises solely from the way this criterion is framed in Australian heritage practice. It does not accommodate the situation where many individuals independently hold the same value'. (Johnston 2023, 247)

However, people perceive their daily environment and everyday heritage most likely as an individual within cultural and social boundaries, not as a group, continuously negotiating identity and attachment (compare Jacobs 2006, 'mindscape'). Individual views on landscapes, places, buildings and objects are what constitute dynamic heritage practice in a daily context. In contrast, not all members of a community share all values at the same level. Interest groups work towards one or a set of aims. Communities are connected through, e.g., locality, profession, tradition, historical interest or local initiatives, for example, neighbourhood planning or stewardships. But communities are not homogeneous, and any approach needs to acknowledge this form of value creation. Individuals can express values that loosely create value communities based on specific landscapes, areas, goals or interests. Sometimes the reasons behind place attachment and connections are also hard to define and express by people themselves. This makes it more challenging for local authorities that want to engage people through local initiatives, such as Local Listing projects.[5] These list applications demand research by the people who want to put local heritage forward which deters people from engaging. The challenge going forward is to find tools and methodologies that allow authorities to collect place attachment data on an individual level and integrate the results of the analysis into the assessment framework for valuation of the historic or everyday landscape.

6.3 Current landscape of place-based, people-centred initiatives

Recent Arts and Humanities Research Council (AHRC) place-based projects acknowledge that the meaning of place-based and people-centred varies. Place, however, can be more tightly defined, as:

> ‘[…] where life courses are shaped, social networks are formed, and the sites of lived and felt experiences. Place is also a geographic location where economic resource is allocated, boundaries are mapped, and data is collected’. (Madgin and Robson 2023, 6)

[5] https://local-heritage-list.org.uk/south-yorkshire Sheffield local listing project

Heritage plays a central role in the process of creating a sense of place and 'is both an input and an output of the process of heritage creation' (Graham, Ashworth and Turnbridge 2000, 4). AHRC- funded place research projects in the UK show the different approaches to express the various forms of place attachment, sense of place and identity, and place making through rootedness (Madgin and Robson 2023). Public involvement and participation have become a focus for development over the last three decades, and the role of experts has been at the centre of discussions. English Heritage campaigns and guidance with titles such as *Knowing your place* (English Heritage 2011a) and *Our Places* (English Heritage 2011b) tried to actively engage people with the heritage on a local basis, but the relevant places were – in the end – still defined by experts. The public were included successfully as stewards, visitors or volunteers but not in actively defining and shaping the values of the cultural landscape. As Smith (2006, 94) points out, the way in which heritage is managed has an influence on public opinion. Schofield (2014) argued that while we still need experts, the role of the public needs to be reinforced. It is the local population who are the experts in their own place and know best what direction change could take in line with locally held values. Schofield's provocative title *Who needs experts?* was critiqued for dismissing the important role experts still play in the process of heritage identification and management and for handing over the field of heritage to an untrained public with no real chance of changing the underpinning policies. It was polemicised that in this process, the whole sector, including academic education, would be at risk. Instead, it was suggested to study up to be more effective (Hølleland and Skrede 2019, 833).

However, so far, several projects have proven that the bottom-up approach is more people- centred and empowering for marginalised groups in society to make their voices heard than a top- down approach, and, while expertise will still be a regulating factor in heritage and landscape management, the role of experts has to change to remain meaningful and inclusive in a changing world (Avrami 2009, 179; Byrne 2008, 15-16; Chitty 2016, 7; Jokilehto 2016, 31; Primdahl and Kristensen 2016; Smith 2006, 4). The role of the public is still seen as a passive participant and consumer. This is, for example, reflected in the language of heritage organisations which is dominated by terms such as 'invited', 'learn', 'share' (Smith 2006, 44). This language contradicts any intention of developing meaningful participation or co-creation. The future challenge will be to design approaches for meaningful integration of public views and social values; also, to create a methodology for a practical, repeatable, on-going evaluation of such data. In view of landscape characterisation[6], for example, there is a sense in which people's perceptions will play an increasingly essential role in the recognition of distinctiveness and quality of places.

6.4 Examples of people-centred approaches to place evaluation and management

Internationally, projects have developed people-centred methods and approaches for the evaluation of people's perceptions. Assessing the personal connection to places that define cultural landscapes has been identified as an essential element of heritage in view of contemporary cultural association, for example in the US (Altman and Low 1992; Low 1987, 2002; Lynch 1960; Manzo and Perkins 2006) in a rural and urban context and Canada (Maguire 2017) associated with national parks, in Italy (Nardi 2014), in the UK (Jones and Leech 2015), in Australia (Byrne 2008a, 2008b; Brown, Raymond and Corcoran 2015; Modesto and Waterton 2020) and in Denmark (Primdahl and Kristensen 2016), the last with a strong focus on heritage practice and landscape perception of marginalised groups and in urban contexts. People's individual perception plays a role in the study of place attachment and has been extensively explored in Poland (Lewicka 2011b). The following section will detail some of these approaches.

Examples of counter mapping can be found in the US with the work of Lynch and his approach to understanding the use of place and attachment within an urban context (Lynch 1960). A similar approach was undertaken by Byrne, mapping the life stories and memories of First Nation People in an Australian context (Byrne 2008a). The approach went beyond and critiqued a one-sided western and nature-focused view on landscapes. Also, Cultural Mapping is an initiative that enables community participation and identification of social values across wider landscapes[7]. Nardi (2014) applied a similar people-centred approach in Italy by using 2D paper maps which were annotated with stories and memories by the local

---

[6] For example, Historic Landscape Characterisation as planning tool for the historic environment. The maps are based on the historic processes that lead to present landscapes and provide the element of time-depth to the assessment process.

[7] https://bangkok.unesco.org/content/cultural-mapping

community. These projects were based on methodologies which provide a good foundation for new approaches using innovative methodologies through advanced GIS and computer capabilities and the evolving field of AI technologies (see also Jones 2017; Jones and Leech 2015).

Common Ground, an environmental and arts organisation, was formed in 1985 by Angela King, Sue Clifford and Roger Deakin out of the environmental movements of the 1970s (Common Ground 2019; Hayden 1995, 63). Their Parish Map project was different from the usual environmentally focussed projects in its focus on local heritage and artistic exploration of local distinctiveness (Perkins 2007, 128). What started with a few artists exploring a sense of place grew into a community-led activity across the country. Parish maps were the visual representation of the essence of a place, the artistic expression of people's attachments and experiences in their neighbourhoods (Common Ground 1996; Perkins 2007, 130). The examples vary from woven panoramas, as produced during the parish maps project, to web-based maps since digital online mapping has become affordable and accessible for the public (Perkins 2007, 127). The artistic expression format made a meaningful integration into frameworks and structures of local authorities for the planning process challenging, if not impossible. The essential development to render such projects integratable was, therefore, the step towards digital technology, which is increasingly finding entry into the fields of heritage and archaeology.

Two projects carried out in Denmark encouraged communities to use landscape characterisation carried out by the communities themselves, based on participant's perception of the environment (Primdahl and Kristensen 2016). This approach allowed the public to inject their views into the planning process and represent on a map the values of the community in the form of 'landscape as a common good' (Primdahl and Kristensen 2016, 229). The advantage of these case studies was the close cooperation between the authorities, specialised experts, academic researchers, other stakeholders, such as landowners, and a group of interested people from the communities. The method proved to influence the planning proposals. However, while the groups continued their work after the end of the research projects, the project was initiated, led and supported by experts and technology which is not usually readily available and accessible to communities (Primdahl and Kristensen 2016, 236). The approach of recreating the landscape character map to include people's views without funding and technical/IT support proved challenging for local authorities with limited resources. This approach has the advantage that it is proactive, not reactive. However, while the mapping process in cooperation with the researchers was successful in representing the values of the communities, one of the case studies revealed that the assessment of natural and cultural characteristics in the landscape proved to be problematic for non-experts (Primdahl and Kristensen 2016, 236), proving the point that experts and local residents need to work together to achieve the best result for increasing the quality of places.

Other projects in the UK, such as *Roots and Futures* by the University of Sheffield (University of Sheffield 2020), *Heritage Lincoln Connect* (City of Lincoln Council 2011), *Know your Place* in Bristol and the English West (Bristol City Council 2021), attempted to capture social values on online web map platforms to evaluate and incorporate the data into the practice of local authorities or to explore how these data can contribute to a holistic view on heritage resources. A further project studying the relationship between museum collections and people during the COVID-19 crisis was launched by Liverpool Museum, providing extraordinary insights into how people's view on the world and their stories changed during the pandemic (National Museum Liverpool 2021). All these projects use digital platforms for the collection and visualisation of the data, which constitutes a development towards meaningful integration opportunities in the planning and assessment structures.

Community-led characterisation by community or focus groups has the potential to empower communities and include their values meaningfully in the planning process (Dalglish 2018, 53-55). However, this approach proved challenging as communities do not always share a coherent identity or are motivated to act purely democratically. Rather, community groups represent the dominant voices in a community rather than the whole community or individual opinions. Furthermore, groups often consist of self-selecting active community members, while others are marginalised leading to 'unrepresentative views' (Dalglish 2018, 56-58). Parts of the community are excluded based on, for example, time restraints, ethnicity, IT illiteracy or anxiety. Despite these difficulties, toolkits have been developed to help communities identify what is important in their neighbourhood. In Scotland, for example, *Talking About our Place Toolkit* (NatureScot 2020) or the Place Standard *How good is our Place?* (Scottish Government 2023) are useful to improve participation in local characterisation and evaluation schemes. Similarly, Robson (2021) developed a methodology to empower communities and qualitative researchers to engage with communities and explore shared values.

Surveys or polls conducted by local governments or organisations to collect public opinions run the risk of consciously or subconsciously influencing or manipulating public opinion. Opinion can be influenced through directed questioning, narrowing the opportunity to express opinions openly. For example, stating that heritage is important and asking the question afterwards or focusing the surveyed group on people who are already engaged with heritage, results in a highly biased survey outcome and brings circularity into the process. Examples are IPSOS Mori surveys formulating statements as questions and asking for the degree of agreement (English Heritage 2000, 2; Byrne, Brayshaw and Ireland 2003, 64; Smith 2006, 94).

According to current views, public participation cannot accommodate every single, individual view of members of a community, and they are not always uncontested and democratic (Dalglish and Leslie 2016, 217). Perkins (2007, 131) shows that even grass roots projects such as the Parish map of Mottram-in-Longdendale revealed tensions amongst communities; agendas were dominated by some groups or parts of the community, while others were excluded. Avrami (2009, 182) suggests, therefore, to focus more on the ‘process’ of ‘negotiating change’ with the public in planning rather than the outcome. In this context, the following section presents a project that developed a methodology to collect and analyse individually held values in places as stories that represent personal connection or place attachment that allows meaningful integration of personal, subjective and individual values associated with places (tangible and intangible factors for connection) into the framework of planning and decision-making.

6.5 Case study: from individually held values to Social Landscape Characterisation

Developing a *Social Landscape Characterisation* (SLC) based on the perception, needs and visions of the local population and visitors, against the background of current practice and previous projects, using bottom-up approaches, was the main focus of a project at the University of York. The used data collection methods comprised social media data (Twitter, now X), online surveys and semi- structured interviews. The resulting datasets were analysed using Artificial Intelligence tools, correlated and visualised in GIS. The resulting digital maps provide opportunities for meaningful integration of the insights into individually held values into assessment frameworks.

While the study areas were located in the UK, the method is applicable anywhere where comparable data exist, in both rural and urban areas. Study areas for this project comprised the Peak District National Park (PDNP) and the city of Sheffield, two spatially closely connected, in fact overlapping, but very different areas in the north of England (for map see Tenzer and Schofield 2024, Figure 2). While the study included the city of Sheffield for the survey approach (Tenzer and Schofield 2024), the focus in this paper is the PDNP, which provided data for all three data collection stages (social media, survey and interviews). The PDNP was the first designated National Park in the UK (in 1951), with wide views, open space and natural beauty. The rich history of human occupation extends from the Palaeolithic, through Roman and medieval times into the historic landscape of post-medieval and contemporary worlds. Archaeological and historical features comprise designated monuments such as burial mounds, henges and stone circles, traces of mining and medieval industries, listed buildings, and conservation areas, as well as a large number of undesignated, locally important heritage assets. Together these historic sites form the environment in which people follow their daily routines, contributing to the distinctiveness of places through their social practices, affording these places values through connection and sense of place, belonging and identity. As elsewhere, communities and places in the PDNP are affected by local and global challenges. Increasing footfall of visitors, community coherence loss through second and holiday homes, wildfires and droughts as results of climate change are just some challenges the National Park has to tackle.

Figure 6-1 shows how the approach taken in this research can be positioned within current assessment strategies, which are (a) expert-led or people-centred (vertical scale) and/or (b) focused on predefined community groups to assess shared values or based on individual stories (horizontal scale). The proposed method is expert-led and implemented but includes people's perceptions through their individual stories. The approach has the potential to form invisible communities and reveal patterns of social values shared by individuals. Such individual values when viewed as single occurrences are anecdotal and subjective; however, when patterns emerge across landscapes such individual values can be understood and interpreted as shared or social values focused across wider landscapes, including everyday and designated heritage.

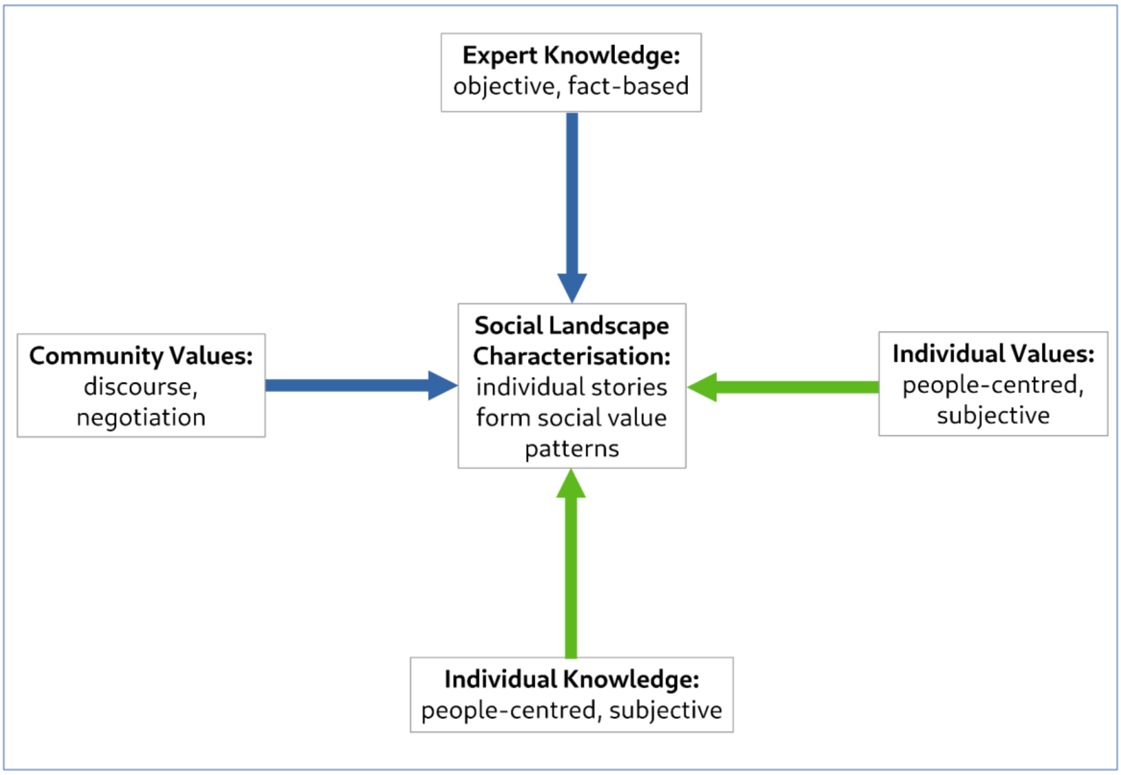

Figure 6-1: Position of Social Landscape Characterisation in the approaches of current assessment strategies. The vertical axis shows the position between expert-led and implemented assessment, which is fact-based, and objective and a people-centred approach based on individual knowledge. The horizontal scale represents the spectrum of community-based values and an approach from the individual story to extract patterns and create invisible communities of shared values based on same meaning or same location.

The methodology consists of three stages: (1) data collection from three different data sources, (2) the application of Artificial Intelligence tools (Natural Language Processing (NLP), Topic Modelling (TM)), after Jones 2021; Jones, Doane and Attbom 2021) and (3) the creation of various outputs for different requirements of potential users, such as community groups or local authorities (for detailed methodology and workflow diagram see Tenzer and Schofield 2024). The project used social media data from people posting about the study areas (Twitter data, now X) (Tenzer 2022), online surveys focused on people living and working in the study areas (Tenzer and Schofield 2024), and semi- structured in-depth interviews with people living and working in the study areas (Tenzer and Schofield 2023). Social media data was based on hashtag searches for the area of the PDNP, extracting

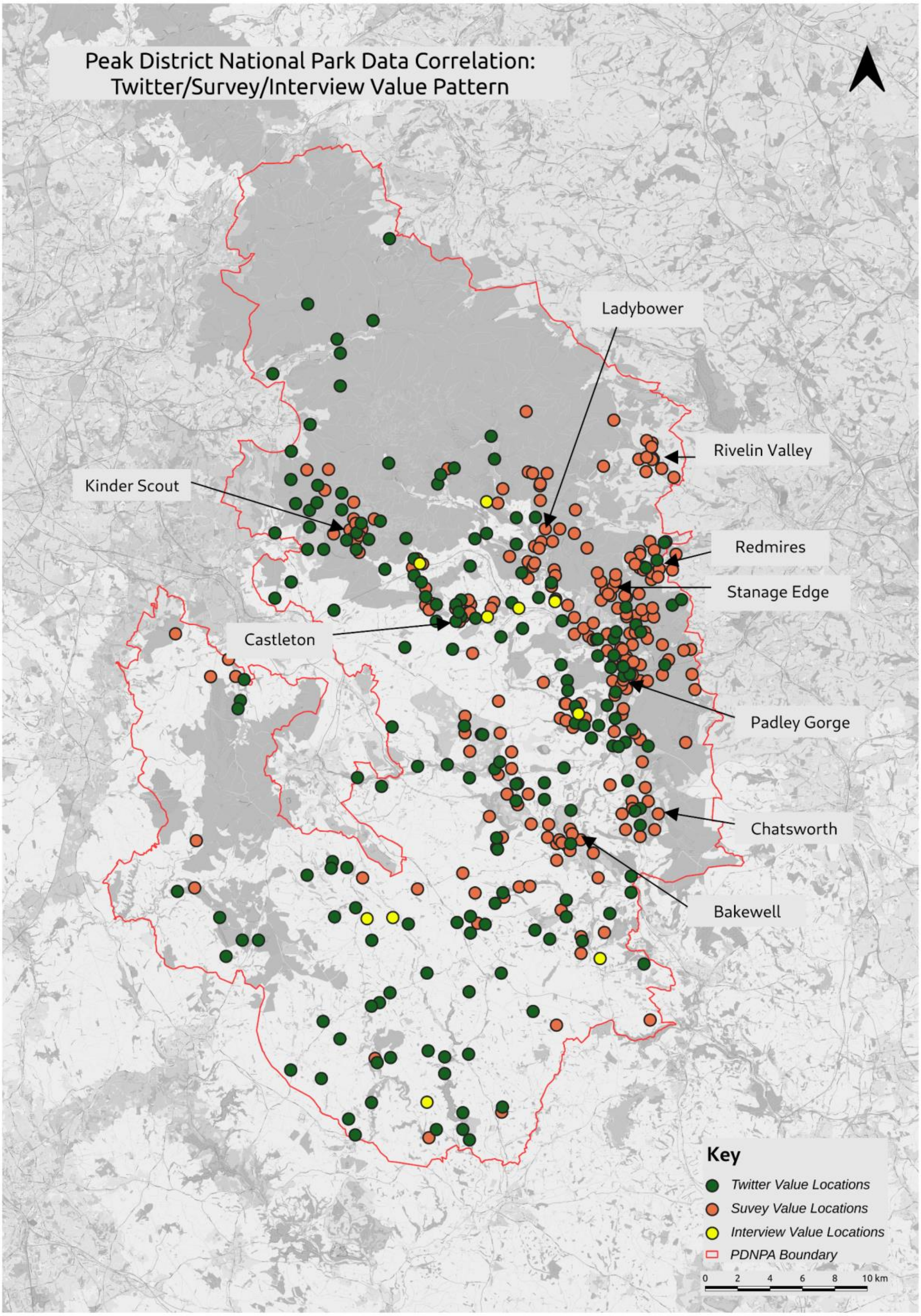

Figure 6-2: Correlation of the results from social media data (green points), surveys (orange points) and interviews (yellow points). The point data symbolises the story and connection of an individual person to a specific place in the PDNP. In this visualisation, the pattern of individual values shared across wider landscapes can provide an insight into areas of high value. However, blank areas symbolise no data not the absence of significance for the people. (map created in QGIS; data contain OS data © crown copyright and database right 2022. Map tiles by Stamen Design, under CC by 3.0. Data by OpenStreetMap, under OdbL).

frequently mentioned locations and sentiments towards those places. The online survey asked participants for their favourite places within the study areas and their stories of personal connections to those places. The interviews were semi-structured to give participants the widest possible freedom to tell their stories. While social media data provided a broader overview of the places dominating the conversation and the emotional connections formed towards them, the survey narrowed the focus and provided insights into the personal stories of connection and perception of local people. A further focus with individual case studies through in-depth interviews gave essential snapshots of personal place attachments and the reasons for the development of a strong sense of place and belonging.

Figure 6-3: Impression of the moorland below Stanage Edge with wide views and remnants of the millstone industry (photo by the author).

Figure 6-2 shows the correlation of the three different data sources in the PDNP study area. Survey and social media analysis provided larger datasets that represented patterns of values in varying levels of convergence. Survey data represent the valued places of locals (orange point data), while social media data (green point data) represent a wider view on valued places from visitors. Hotspots, representing a stronger connection of local people, are apparent in, e.g., the reservoirs of Ladybower and Redmires and Rivelin Valley, as ideal local recreation areas. Similarly, Stanage Edge as a natural landmark featuring remnants of the industrial past in the form of large millstones (Figure 6-3) and the market

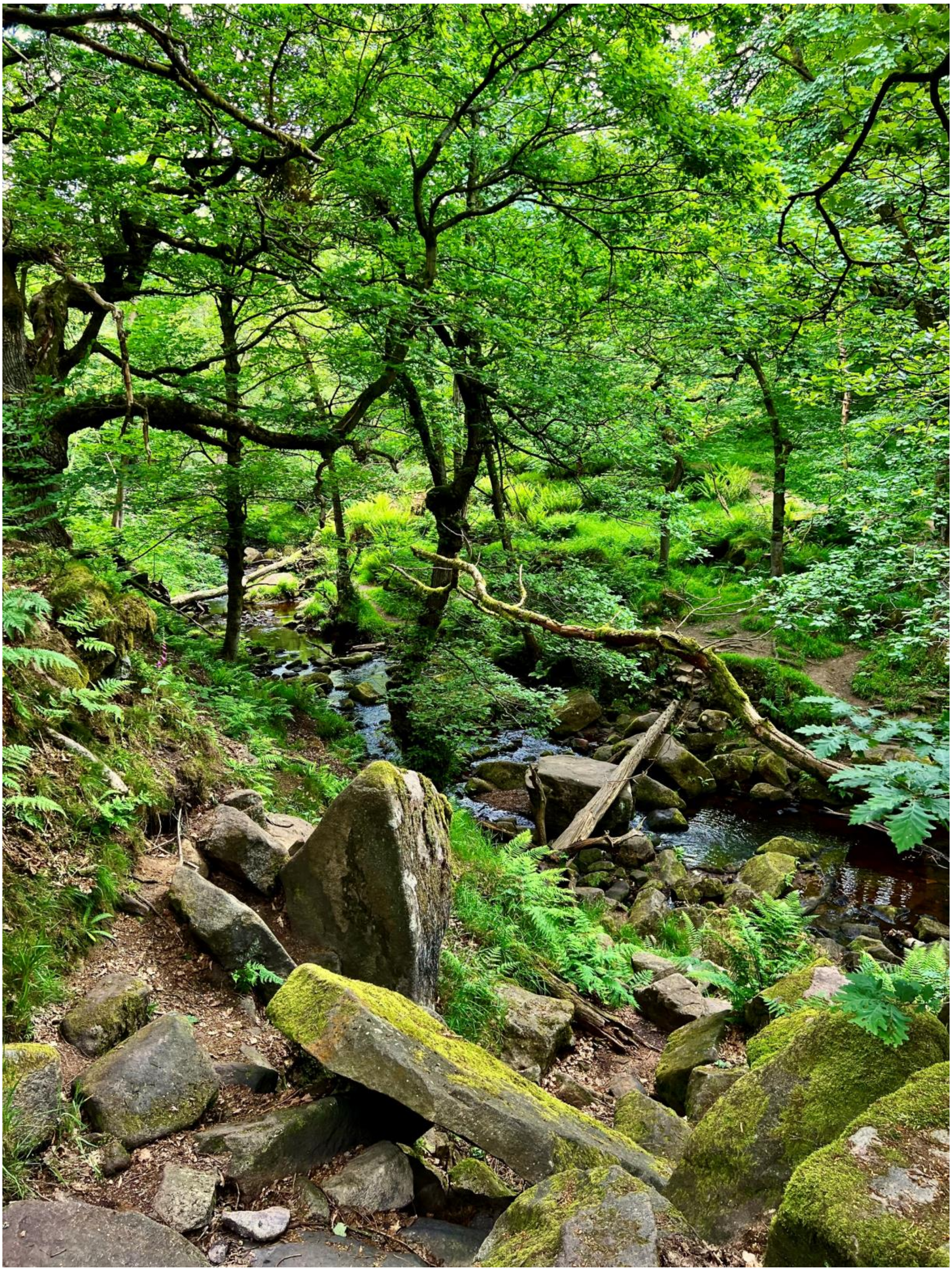

Figure 6-4: Padley Gorge, a rare example of a temperate rainforest in the UK, a place valued by visitors and locals for the natural qualities and as recreation space. While the place seems natural, wild, and untamed human traces can be tracked across the landscape, e.g. drystone walls and remnants of the millstone industry (photo by the author).

town of Bakewell are seemingly more valued by local residents of the PDNP and Sheffield than by outsiders. In contrast, the south of the PDNP, also known as White Peak, with a gently undulating landscape, deep gorges and caverns, was more often mentioned by social media users tweeting about the PDNP. Both data sources revealed high values for, e.g. Padley Gorge (Figure 6-4), one of the UK's rare temperate rainforests, valued for its qualities for recreation and natural beauty. A high significance for social media users and survey participants was notable at Kinder Scout, the highest point in the Peak District with wide views, moorland and waterfalls. The yellow point data on the map (Figure 6-2) represents the interview locations (for more details of the methods and results of this approach, see Tenzer and Schofield 2023). This in-depth exploration of people's deep connection and rootedness in the landscape focused on the reasons behind a strong place attachment. The participants were partly located in areas that were identified as highly valued by social media users and survey participants (northern part of the PDNP) and areas of less dense data (southern part of the PDNP). The reasons for connection were dominated by the

Figure 6-5: Impression of the valued landscapes in the PDNP and Sheffield: Rivelin Valley and the wheels along the watercourse. The place affords recreation and historical connection and represents the qualities of nature/culture heritage (photo by the author).

aesthetic qualities of the landscape, the rich historical past and personal memories, and life histories in the particular places.

The results revealed two distinctive patterns of invisible or hidden value communities, based on single, individually held values (for detailed results see Tenzer 2022; Tenzer and Schofield 2023; 2024). While current evaluation processes define a community in advance of research or assessments and negotiate dominant values, this approach allowed such value communities to arise and form naturally based on two factors: first, people who connect to the same places for different reasons – location-based communities; and second, people who value wider areas for the same reasons – value-based communities.

As an example for the first category, participants in the survey favoured the 'natural' qualities of the landscapes in the PDNP in various spots (see categorisation of the data in Tenzer and Schofield 2024, Fig. 8; also Bell 2005). Reasons for this connection were often associated with the restrictions of the COVID-19 pandemic and subsequent changes of behaviour and the need for relaxation and recreation in an increasingly busy world. Such seemingly natural and untamed, wild places are the result of human impact on the environment but treated as natural heritage, creating the dichotomy of nature/culture in the discussion of heritage and landscape management (see Byrne and Ween 2015; Harrison 2015). This dichotomy did not exist in the minds of the research participants and reasons for attachment overlapped, spanning from aesthetic values, family and local history value, the need for recreation and exercise to spaces of mental well-being and solitude or community and cultural programmes.

The study also showed that people connect and appreciate historical places that are officially designated as heritage, such as the Grade I listed stately home of Chatsworth House in a Grade I listed Park and Garden (Figure 6-6). Other locations, e.g. the small town of Castleton, are favoured mainly by locals for associated traditions, such as the Castleton Garland and Well Dressing. Figure 6-5 gives an impression of Rivelin Valley, one of the river valleys leading from the PDNP into Sheffield and providing green space for recreation as well as a connection to the industrial history of the city. The second category or the accumulation of different values in specific places was present in the data and visualised as hotspot maps (Tenzer 2022) and category maps (Tenzer and Schofield 2024). Important to note is that blank areas on the map represent a lack of data and not the absence of values or of their significance for people.

How an inclusion of social values in the planning and decision-making process might look is shown in Figure 6-7. The proposed structure contains maps with different information from various sources, e.g. Historic Landscape Character data, results from social media research or surveys, with variation in complexity of represented data achieves with varying levels of application in the planning process and potential for engaging resources for outreach and public engagement. The results prove that the collection, analysis and meaningful integration of individually held social values can be achieved, using time-efficient methods, such as TM and online surveys, and digital visualisations in GIS that can be integrated in existing tools of local authorities.

Figure 6-6: The stately home of the Duke and Duchess of Devonshire, a Grade I listed Chatsworth House set in a registered Park and Garden, are officially designated heritage assets and equally valued by local people and visitors, albeit for different reasons. While the designation is based on the historical value of the country house and the designed parkland, people also value the place for the recreational qualities and traditional events (photo by the author).

Figure 6-7: Framework for inclusive heritage mapping with decreasing complexity and increasing potential for engagement and outreach from top to bottom. Information on each map varies and is gathered from different sources showing HLC data, results from social media research, survey results and other. The combination of information allows to create a comprehensive data structure of landscape information for specific areas.

6.6 Discussion: the study in context

The methodology outlined in this paper builds on the shift in society which started as the cultural turn in the 1960s and was conceptualised in charters and conventions concerning the cultural and natural heritage, such as the Faro Convention (Council of Europe 2005) and the Burra Charter (ICOMOS 2013). In response on a national level, social values were added to the traditional canon of heritage values in the official assessment strategies as shown in the example of the Australian and UK guidelines (Australian Heritage Council 2009; English Heritage 2008). However, the realisation of the aim to include people and local knowledge meaningfully into heritage practice and management has not yet been fully realised. The reasons for this are the continuing adherence of professionals and experts to the AHD

(Smith 2006) and the challenges of capturing and including social or individually held values in the framework of assessment strategies (Dalglish and Leslie 2016, 217). Sustainable, inclusive and transparent planning needs to acknowledge the people-place connection and better understand the reasons behind place attachment to provide better quality places in future. This need is fully acknowledged in the AHRC Place Matters project: 'Understanding place as somewhere with lived and felt as well as geographic and economic dimensions is crucial to the pursuit of better outcomes for people and place' (Madgin and Robson 2023, 6). Internationally, several initiatives are working towards the goal of meaningful integration of people-centred perspectives into planning for better places in rural and urban landscapes (Jones 2017; Madgin and Robson 2023; Nardi 2014; Primdahl and Kristensen 2016). However, projects have a slow uptake as IT literacy and access to the online resources can be challenging for the wider public. Also, research projects often rely on licenced software and highly skilled digital researchers. Such support usually ceases after project funding ends.

This research shows that individually held values can be collected, analysed and visualised in a time efficient way. Instead of predefining groups or communities, participants were allowed to tell the stories that connect them to their favourite places in a given study area. The resulting patterns, formed based on the same values across the landscape or different values in the same place, revealed hidden value communities. This provided the opportunity to infer and categorise the reasons behind a strong place attachment from personal stories, which allowed the creation of value categories based on the language and themes dominant in the empirical data, as opposed to a predefined and expert-led top-down process. This people-centred approach will offer new ways to explore people's views on natural and cultural landscapes in a working and living environment.

The presented case study provides an example for the innovative application of AI methods to qualitative data in a landscape management context. NLP and TM were used for the analysis of unstructured text documents following the principles of Grounded Theory (Charmaz 2006) to achieve a first bias-reduced insight into the latent themes within the data. This allowed themes to emerge from the empirical data which were not anticipated in the outset of the project. A subsequent manual analysis revealed shortcomings, ethical implications and advantages of the method (for details see Tenzer and Schofield 2024). A refined study could include and focus on marginalised groups and explore pathways to engage people that were not reached in this study. A further project would also have the potential to use Machine Learning and Deep Learning methods to automatically categorise new data by unsupervised/self-taught learning. The field of AI deployment in heritage and landscape studies offers a wide range of opportunities in the future.

The trans- and cross-disciplinary development of digital technologies and AI have yet to find a way forward in the field of landscape and heritage management. However, data scientists, digital archaeologists and heritage professionals are starting to collaborate for a joint approach to the application of digital methods in archaeology and cultural heritage management. Such collaboration will in future open new ways and give new insights into new and existing data sets to explore and use archaeological and historical data for the creation of resilient and coherent communities and have a positive impact on place-making and care for the environment in the face of global challenges.

6.7 Conclusion

This paper presents and gives context to a novel methodology that allows capturing and visualising individually held values and analysing reasons behind strong connections to places. The results show that it is possible to integrate and categorise individually held values and visualise these as coherent patterns of shared social values across wider landscapes.

Although this case study is focused on a small area in the north of England, the methodology can be scaled and applied much more broadly. Current advances in internet connectivity, analysis tools, e.g. AI tools, and technologies for interactive collaboration and visualisation, e.g. online surveys and GIS, have the potential to realise the ambitions of the cultural turn for a meaningful discourse between planning authorities and community groups as well as individual people that benefit from positive change and preservation of living landscapes. Data sets of this Social Landscape Characterisation offer a vital background for proactive planning and decision-making within local authorities while integrating people's individual values meaningfully into the heritage and landscape management strategies.

The study of story-based, individually held values has the potential to generate practical applications with an effective integration of people's views into the framework of planning policies. It shows that individual opinions based on local people's personal stories, memories and traditions can be meaningfully and efficiently integrated into the official frameworks of planning and decision-making. Reinforcing the bond between people and place through a bottom-up approach within local planning and thereby generating appreciation of the everyday places and a wish to care for this environment, might help to tackle the most pressing problems of the present but also of future generations and foster cooperation and communication between planning authorities and the people for whom inclusive and transparent planning strategies provide better places to live.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This project was carried out as a part of an AHRC/UKRI WRoCAH-funded PhD project at the University of York Grant reference number: AH/R012733/1.

Ethics approval

The study was approved by the ethics committee of the Department of Archaeology, University of York, UK. All participants of the surveys and interviews have provided informed consent. Social media research adhered to the terms and conditions of the respective social media platforms.

Supplementary material and additional information

There is no supplementary material available for this publication. The references can be found at the end of this thesis. The ethical approval as well as accompanying documentation is available at: **D. Appendices for Chapter 8: Appendices 6 and 7**, Consent Forms: **D. Appendices for Chapter 8: Appendices 3 and 5**, Information Sheet: **D. Appendices for Chapter 8: Appendix 4**.

Chapter 7:

Debating AI in Archaeology: Applications, Implications, and Ethical Considerations

Published as:

Tenzer, M., Pistilli, G., Brandsen, A. and Shenfield, A. (2024). Debating AI in Archaeology: applications, implications, and ethical considerations. Internet Archaeology 67. At: https://doi.org/10.11141/ia.67.8

Abstract

Artificial Intelligence (AI) is not a recent development. However, with increasing computational capabilities, AI has developed into Natural Language Processing and Machine Learning, technologies particularly good at detecting correlations and patterns, and categorising, predicting, or extracting information. Within archaeology, AI can process big data accumulated over decades of research and deposited in archives. By combining these capabilities, AI offers new insights and exciting opportunities to create knowledge from archaeological archives for contemporary and future research. However, ethical implications and human costs are not yet fully understood. Therefore, we question whether AI in archaeology is a blessing or a curse?

7.1 Introduction

Although it might seem so, given the current AI hype around Large Language Models (LLMs) and generative AI models for content generation (such as ChatGPT), Artificial Intelligence is not a recent development. Deployment of the technology in the fields of archaeology and heritage studies with both object and remote sensing applications has been widely documented (Bickler 2021). With recent developments and advances of AI tools in the field of text-based analysis, this will be the primary focus of this paper.

The term Artificial Intelligence was coined in 1956 (Russell and Norvig 2016) describing a hypothetical computer technology developed by Alan Turing (Turing 1950). Following the first AI hype of the 1950s and 60s – over-promising the capabilities of AI technology but under-performing due to the lack of computational power – AI research was interrupted by the AI winter of the 1970s and early 1980s. However, after 60 years of exponential growth, AI tools have now entered the mainstream. Examples include chess computers, recommendation systems, and spam filters. Other applications are now leveraging the recent developments in LLMs, for example, the Google search function, instant translations, and closed captioning.

Increasing computational capabilities enabled the development of Machine Learning (ML) and Neural Networks (NN). In particular, Deep Learning with its ability to learn features of interest in parallel, e.g., the attention mechanism in LLMs, pushed AI capabilities. These systems are particularly good at detecting correlations and patterns, and can categorise, predict, or extract data in the context of natural language processing. LLMs, such as Google's BARD, OpenAI's ChatGPT, or Meta's LLaMA now form the basis of a new generation of Open Source LLMs, such as Open Assistant (Köpf et al. 2023). These tools can learn and draw from extensive datasets that are based on the wide knowledge of the Internet, including data from, for example, Wikipedia, GitHub, and Google data search.

Following an early adoption of AI technologies in archaeology for objects and remote sensing applications (Bickler 2021; Argyrou and Agapiou 2022), NLP, ML and DL are now being used for processing vast amounts of data accumulated over decades of research. This knowledge deposited in archives and grey literature can be efficiently analysed, structured, and disseminated using AI technologies – an approach that offers new insights and knowledge extraction from archaeological archives as never before.

However, while the deployment of AI technologies based on LLMs are capable of processing big data in archaeology and other fields, their application also has ethical implications. The lack of transparency of content and quality of the training data has been shown to reinforce social inequalities, misinformation, privacy issues, racial discrimination, risk to natural resources, and human workforce exploitation. Some of these are the same concerns across the discipline of archaeology and CHM, specifically regarding sensibilities around privacy, bias, and model creation in the context of policy and decision-making.

In this paper, we focus on archaeology as part of that wider debate and present examples of successful AI applications in archaeology with text-based analysis as primary focus. We then provide insight into the ethical implications associated with AI before discussing the implications and applications of AI in a safe, sustainable, and socially just way in future. Finally, we want to open the discussion to the question if AI is a blessing or a curse for the discipline.

7.2 Applications of AI in archaeology and CHM

Archaeologists have a long tradition of adopting, adapting, and introducing technologies from other disciplines. For example, the pantograph preceded digital photography or survey methods (Novaković 2018) while Lidar has proved useful for detecting sites particularly across difficult terrain (Cohen, Klassen and Evans 2020). AI image recognition techniques were introduced in archaeology for remote sensing (Verschoof-van der Vaart et al. 2020) and object recognition (Anichini et al. 2021).

However, adopting AI technology for text analysis is more challenging. Language is complex with ambiguities and hidden meaning beyond the pure text structure. Yet, NLP has immensely benefited from the integration of LLMs. Machine and Deep Learning have been applied, for example, to archaeological prediction and detection (Resler et al. 2021) and CNN to translate cuneiform tablets of old Sumerian and Akkadian languages (Gutherz et al. 2023). Generative AI is helping to recreate the landscapes of the past for more immersive research of the past (Cobb 2023). Big data has been successfully linked in the project 'Unpathe'd Waters (Eagles 2022).

A current cultural heritage project applied NLP and in particular Topic Modelling (TM) and ML to explore the values attributed by people to familiar cultural landscapes

(Tenzer 2022; Tenzer and Schofield 2024). Social media data, online surveys, and interviews provided sufficiently large datasets to infer heritage values from a "bottom-up" or people-centred perspective. TM allows the identification of patterns as themes latent in or emerging from the data, which guarantees an assumption-free approach to empirical data.

AI can also deal with the data deluge being experienced by archaeologists (Bevan 2015). The AGNES project facilitates large-scale synthesising research in The Netherlands, by integrating ML into a search engine which aims to index all the texts about archaeology in the region, some 200,000 documents. Specifically, it uses Named Entity Recognition to automatically detect all time periods, artefacts, and place names, which can then be used in search queries. This allows for more exhaustive and more precise searches, and in a case study on Early Medieval cremations, led to 30% more cremations being found in the literature than were previously known (Brandsen and Lippok 2021).

As well as AI-assisted search and TM, recent advances in the application of LLMs in NLP have shown promise in the identification of personally identifiable information (PII) and potential copyright infringements in digital publishing of archival data from modern historical periods. Legislative requirements (including those imposed by the EU's General Data Protection Regulations and extensions of copyright terms) mean that publishers of historical and heritage archives currently need to spend significant amounts of time and manual effort on ensuring compliance in these fields. Supporting publishing and editorial teams in this process has significant benefits in terms of both the amount of material that can be digitised and published and in catching cases of infringing content that might have otherwise been missed.

However, as useful as the technology seems to be it comes with a human and environmental cost. In the next section, we will present the challenges and risks of AI deployment from an ethical and environmental view as a counterbalance to the advantages and opportunities.

7.3 Ethical considerations – exclusion, limitation, bias

The latest AI advancements have given rise to several ethical considerations that warrant thorough examination. In particular, concerns have been raised regarding the transparency of the content and quality of the training data used in AI applications (Bender et al. 2021).

These factors have been shown to perpetuate social inequalities (Casilli 2019), propagate misinformation (Wilner 2018), and compromise privacy (Véliz 2021). Furthermore, the use of AI technologies has been linked to instances of racial discrimination (Raji et al. 2020), the endangerment of natural resources, and the exploitation of human labour (Crawford 2021).

Within the discipline, concerns surrounding privacy, bias, and model creation, are critical for formulating policies and decision-making. For instance, AI algorithms in analysing archaeological data could inadvertently lead to biased interpretations of historical events or the reinforcement of existing power structures if the models used are not designed with these ethical considerations in mind. Specifically, potential harms of fostering a linguistic monoculture, unintentionally strengthening existing power structures, and becoming a monocultural value carrier (Johnson et al. 2022; Pistilli 2022). Archaeology being also about understanding human history through material remains, language becomes a key component of cultural heritage and identity. If archaeological narratives are dominated by a single language or cultural perspective, this can lead to a skewed understanding of the past, privileging certain histories over others.

Also, there is a need for explainability and transparency in the approach to data collection in qualitative research. As shown in the heritage case study, AI can help analysing vast amounts of social media data or survey responses. However, generating models based on such data can introduce or reinforce biases, for example, excluding already marginalised groups. Shaping policies on models trained on such data would introduce these societal inequalities into systems of governance. The public also needs to have the option to opt-out with regard to data privacy, particularly in the case of vast data sets that are scraped or mined from the internet for training purposes.

While AI has the potential to analyse vast amounts of data and is particularly good at pattern detection (e.g., Casini et al. 2023), the technology has the potential to replace human volunteers in citizen science projects (Ponti and Seredko 2022). This can lead to a decrease of inclusive and engaging projects within archaeology. Excluding the public from the process of data collection and knowledge creation and instead reducing participation to the final product of archaeological investigations can lead to an alienation of archaeology.

Finally, garbage in, garbage out and black box effects carry the risk of creating new content from already flawed data and in an opaque process (Huggett 2021). Kansteiner

(2022) and Clavert and Gensburger (2023) warn about the risk of using ChatGPT to reshape historical narratives: 'If we think that the stories and images we consume influence our memories, identities, and future behaviour, we should be very wary about letting AI craft our future entertainment on the basis of our morally and politically deeply flawed cultural heritage' (Kansteiner 2022, 124). Similarly, the GenAI technology will take realities of cultural heritage into a new dimension with challenges for authenticity and speculative interpretation in a new era of knowledge production and presentation (Spennemann 2023). A similar effect can be expected in the analysis of large archaeological datasets, shaping a narrative of the past based on weights in hidden layers (Cobb 2023).

Four key messages around ethical considerations result from these observations:

(1) The issue of biases emerging from the data used for training AI models is serious. Therefore, it is crucial to ensure data are as representative as possible. Researchers across the discipline of archaeology and CHM should work closely with data scientists and social scientists to design representative sampling strategies and data gathering methods, and to develop protocols for assessing and correcting for bias in datasets.

(2) The intersection of data science, philosophy, and archaeology suggests the advent of a new kind of archaeological specialism. Within this area of practice, archaeologists will need to understand the nuances of AI and Machine Learning and be well-versed in ethical considerations. Furthermore, users of the new technology have to understand the agency and autonomy of the new technology. Huggett (2021, 428) argues that "in some cases the system can appear to replace human expertise".

(3) The use of AI in shaping historical narratives is controversial. While AI has the potential to analyse large datasets and reveal patterns not always discernible to human eyes, it also carries the risk of propagating flawed interpretations of the past, particularly if the underlying data are biased. Therefore, stringent checks will be needed on the application of AI in this context. This includes the implementation of explainable AI (XAI) techniques to make the decision-making processes of these systems understandable to humans. However, the implementation of XAI techniques - even in simple application domains - is challenging. Two contrasting XAI philosophies exist (Barredo Arrieta et al. 2020) - 1) designing inherently interpretable

AI/ML systems, and 2) applying post-hoc explainability models (such as SHAP (Lundberg and Lee 2017)) to try and explain decisions made by AI models. A key disadvantage of inherently interpretable AI models is that it limits the power and complexity of such approaches - particularly in leveraging the latest generations of generative AI systems; however, criticism has been levelled at post-hoc methods regarding how closely their explanations relate to the decisions made by AI algorithms.

(4) Ethical guidelines for AI applications in archaeology and heritage practice need to be drafted and widely adopted to prevent misses and to promote the responsible use of these powerful technologies. However, crafting ethical guidelines for AI use in archaeology requires a balance between preventing misuse and adapting to the varied legal and practical contexts of global research environments. Discussions at the World Archaeological Congress (WAC 2023) and studies on remote sensing practices (Fisher et al. 2021) stress the challenge of developing standards that accommodate the distinct local regulations and the particularities of conducting research across different cultures and regions. Nevertheless, Davis (2020, 1) argues, that a high level of automation based on algorithms has the potential to create 'consistent definitions which permit reproducible research designs', which shows the advantages of automation for compatibility and reproducibility of data.

7.4 Discussion

Recent developments and the rapid adoption of AI technology into archaeology and heritage practice, as presented in this paper, show the importance of a debate around ethical implications and sustainable applications of AI. To enable the discourse, we have presented the advantages and capabilities of the applications, which allow more time and resource efficient workflows (Tenzer 2022; Tenzer and Schofield 2024), and enable the analysis and reuse of 'big data' accumulated over decades of archaeological investigations lying dormant in archives and grey literature (Brandsen and Lippok 2021). Furthermore, we provide different views on the implications of AI applications from archaeology, heritage studies, data science and philosophy, showing inherent challenges regarding limitation, bias, and social impact (Bender et al. 2021; Casilli 2019; Crawford 2021; Véliz 2021).

Interdisciplinary/cross-disciplinary research and collaboration will be necessary in the near future to apply this technology to a wide variety of disciplines. Collaboration between data science, sociology, philosophy, and archaeology is becoming increasingly important. Understanding how AI technology can influence epistemology and hermeneutics has to focus the discussion on the agency and cognitive artefacts of the technology in view of the output (Huggett 2021, 421). University courses bridging the complex knowledge of the various disciplines will be increasingly necessary. The projects presented here and the collaboration of the authors of this paper exemplify how cooperation can work to foster mutually beneficial collaboration.

Furthermore, the discipline needs to understand how AI deployment will impact on future employment for archaeologists and the changing work environment. What are the prospects for future archaeologists as a professional and academic career? Do we need to become computer scientists ourselves, and teach this to our students? Ultimately, will AI replace archaeologists? Harari (2017) argues that there is 'only a 0.7% chance'. However, it can replace the monotonous tasks of daily work, and carry out the large-scale analyses that precede archaeological work. However, the technology is evolving with increasing speed and predictions of future impact on the profession, especially after the pandemic, are difficult going forward.

AI deployment in the discipline needs to run alongside the development of strategies and best practice guidelines safeguarding the responsible, fair, and sustainable use of this new technology. Exploitation of human and natural resources with a cost for the environment needs to be highlighted and potential risks to reinforce social inequality must be considered.

Archaeology and CHM scholars are well equipped to study and deal with these societal effects of AI, looking at large scale influences on society for decades, and having the theories, methods, and background for these analyses. But to do so, they first need to understand the AI methods and their implications.

7.5 Conclusion

In post-phenomenological ontology, humans are experiencing the world with and through technology (Gattiglia 2022; Ihde 2009). While we are at a point where machines not only assist humans (first machine revolution) but replace humans in the production or creative workflow (second machine revolution), we need to reorientate and redefine objectives. AI is here to stay, and the question will be how to use it responsibly and sustainably.

This means alignment: where does the technology work towards humanities values and goals and where are the dangers and risks of losing control, and therefore the benefits for society and humanity as a whole; not for the benefit of a few, but for the improvement of the environment, health, and society of the many?

Where does the development go from here? How can AI shape the future of the past – increasing our understanding of the past, using the vast amount of data from archaeology and history to create material that promotes and conveys this knowledge? Where does the future of the discipline lie regarding cooperation and education? We are at a point where archaeology and heritage practice cannot only benefit from these technological developments and advances but must also contribute to the ethical and practical discussion of AI in human culture and societies. Coming back to the initial question if AI in archaeology and CHM is a blessing or a curse, we provided examples of advantages and beneficial applications of the technology, but also highlighted challenges that need to be resolved before AI can be used safely and democratically. The debate is wide open.

Funding statement

This project is part of an AHRC/UKRI WRoCAH-funded PhD project. Grant reference number: AH/R012733/1.

Additional Information

The co-authorship of Giada Pistilli is acknowledged for her contributions to conceptualisation, writing (a paragraph on *Ethical Considerations* of the original draft), and editing (10% share).

The co-authorship of Alex Brandsen is acknowledged for his contributions to writing (a section in *Applications of AI in archaeology and heritage studies* and a case study in the original draft) (8% share).

The co-authorship of Alex Shenfield is acknowledged for his contributions to writing (a section in *Ethical Considerations*) and editing (5% share).

Chapter 8:

General Discussion and Conclusion

8.1 General discussion

8.1.1 Introduction

This chapter will assess the outcomes of this research against the aims and objectives set out at the start of the project (see **Chapter 1**). The main findings were published as journal papers, which are included as **Chapters 3 to 6**. While each publication focused on one specific data source and the AI techniques used and developed to answer specific questions on the datasets, the following discussion will compare and correlate the findings of the separate parts and present the connection and correlation between data sources and overlapping techniques applied across the data sources. The main findings discussed in detail in this chapter include the bottom-up approach of the data collection, a distinction between social and individually held values, the cross-fertilisation of methods used and developed for one particular data source and subsequently applied to different data sources, the correlation of values across the study areas, as well as the categorisation of individually held values and their visualisation as patterns across the landscape.

The discussion will include limitations, ethical implications, and bias encountered in every step of the process, with a more research specific view than the general considerations in archaeology and cultural heritage studies described in the AI ethics publication included in **Chapter 7**.

In addition, a section on challenges and successes experienced during the project is included and provides a personal reflection on the approach and related difficult and beneficial aspects of this research. Where the objectives could not be achieved as initially anticipated, the reasons are discussed in detail, for example, where HLC principles were not fully compatible with the principles of the SLC tool developed.

Finally, the conclusion summarises the key findings and provides an outlook on further research which could build on and further develop the methods and approaches applied here. Applications of the method to other areas and target groups, including reviews on an annual basis could contribute to its development and to a better understanding of social values and place attachment through a larger dataset and more opportunities for comparison.

8.2 Discussion in relation to aims and objectives

This thesis aimed to further the understanding of individually held values and create a landscape characterisation based on the resulting picture of social values across historic landscapes for an inclusive and transparent heritage and landscape management. The common approach to landscape characterisation for planning and management purposes is based on expert knowledge and assessment to facilitate change and development and, at the same time, safeguard the historical and archaeological assets in the present for future generations. Initiatives and projects to explore the meaning of elements of the historic environment from a people's perspective usually resulted in artistic expressions, which could not be integrated into local and national assessment frameworks. Also, approaches and initiatives detailed and discussed in **Chapter 1.2** showed the dominant focus on either pre-defined communities or groups or site-based approaches. While official frameworks acknowledge and integrate social values to a degree, neither academic research nor local planning recognise the individual perception of people, as opposed to group or community opinions. There is also no current method to assess individually held values on a landscape-scale to visualise value communities formed by individuals (see **Chapter 1.3**). This thesis proposes *Social Landscape Characterisation* (SLC) to provide such a method and enable the assessment of individually held social values across wider landscapes. This approach adds an innovative method and complements and advances the tools currently available to practitioners and heritage professionals.

International charters and agendas have called for increased engagement and active participation of communities and individuals in relation to heritage and cultural landscapes. To address the shortcomings and challenges related to this and provide a methodology, this research used Artificial Intelligence techniques to assess people's perceptions, needs, and visions individually to establish an innovative routine to collect and analyse individually held values. This routine was also used to visualise and categorise evolving value patterns of 'hidden' or 'invisible' value communities across the rural and urban landscapes of the two study areas, the Peak District National Park and the City of Sheffield. The main aim of the research was achieved by creating a scalable, flexible, and dynamic tool that captures social values and provides a dataset that can be used in conjunction with other datasets for the assessment of natural and cultural landscapes. The *Social Landscape Characterisation* tool allows the integration of people’s perceptions in such assessments, which was initially anticipated as an extension of HLC by its creators but never fully realised.

In the following discussion, I will address the objectives outlined in **Chapter 1** in detail. The objectives set for this thesis were as follows:

> 1: How can people’s individually held values and reasons behind these values (‘soft’ or subjective data) be collected, allowing a categorisation based on latent themes within the data and analysed, using freely available and open-source software and code?
>
> 2: Can Historic Landscape Characterisation or its key principles be adopted and adapted to accommodate people’s perceptions and opinions on their living and working landscapes?
>
> 3: How can social values be visually represented to create outputs for assessment frameworks within the planning and decision-making process and, at the same time, provide opportunities for developing engaging resources to increase participation for inclusive, transparent, and socially sustainable heritage and landscape management? The aim of this objective is to produce a guideline or methodology that can find practical applications in real-world scenarios.

8.3 Objective 1

8.3.1 Restrictions and principles

To achieve Objective 1, data collection exclusively targeted individuals as participants in the data collection for all methods applied. Due to restrictions during the COVID-19 pandemic, the development of data collection methods was limited to remote techniques, such as Zoom interviews (later conducted in-person when restrictions were lifted), online surveys and social media data. This was not considered a disadvantage, although the reliance on computer-based methods and social media users introduced some limitations and biases (see **Chapter 8.7: Ethical implications, limitations, and bias**).

Nevertheless, the objective was achieved by collecting and categorising individually held values based on personal tweets and stories of people narrating their connection to the working and living landscapes of their favourite special and everyday places in survey and interviews. Latent themes within the data were extracted using Topic Modelling (see **Chapters 3** to **5**, for the lab book see **D. Appendices for Chapter 8: Appendix 1a and 1b**).

In order to achieve an important principle of this research, which is to use only freely available and, ideally, open-source software, I restricted the software used in this project to: Qualtrics survey software, which is freely available as a basic package, and Google Maps, for the creation of a map-based survey tool[8]. Furthermore, I used open-source R programming language in RStudio for coding, NLP and Topic Modelling open-source codes for the analysis and QGIS, an open-source GIS[9], for spatial analysis and visualisation. The data sources were also chosen favouring open access options, which provided valuable base data for this project, for example, GIS datasets provided by the Archaeology Data Service (ADS), Historic England or the Ordnance Survey (OS). This methodological toolset allows the application of the method without a financial commitment to licensed software, meaning it can be applied by individuals, communities, and groups with no funding or local communities with low budgets.

The following sections detail how the objective's components of individual people's data collection and categorisation were achieved.

8.3.2 Capturing stories in the landscape – the bottom-up approach

To gain a broad as well as a deep understanding of reasons for people's connection to places, I used three methods for the data collection: social media data (**Chapter 3**), online surveys (**Chapter 4**) and in-depth interviews (**Chapter 5**). The approach with an open format of the questions, with free text entry for surveys and semi-structured interview questions, allowed the study participants to identify elements of the landscape that matter most to them freely and without directing their responses. The aim of this approach was to give people the greatest possible freedom to tell their stories about favourite places and their

[8] Licensed alternatives for the survey software would be ArcGIS or Maptionnaire (https://www.maptionnaire.com).

[9] Where research institutions or authorities hold a license, other GIS systems can be substituted to better integrate the this into existing workflows.

personal connections to the places independently of the researcher's predefined notion of what constitutes heritage, a historic environment, or natural and cultural landscapes in a wider sense. Participants were encouraged to provide their personal stories and photographs, which would provide an insight into the reasons for the attachment and rootedness in place, creating a sense of place, belonging and identity. There were no limitations, and people were allowed to provide examples of all forms of personal or everyday heritage – not just the nationally important and officially designated sites, monuments, and buildings as elements of the cultural landscape. The notion that everything created in the past is in some form heritage shaped the understanding of everyday heritage in living and working landscapes[10] [11]. The basic assumption of this research is that every individual, as the smallest part of groups, communities, or societies, has an invested interest in this heritage to shape the places they live in in a way that accommodates their needs, vision, and aspirations and contributes to the aspirations of wider society. Elements of the landscape are imbued with meaning, creating a value, which is defined as social value and incorporates tangible and intangible aspects of the environment that roots people in places and creates attachment (Jones 2017; Tuan 1990, 1980).

This bottom-up approach aimed to explore how substantial the gap between experts' assessment and people's perception is and to gauge the opportunities for meaningfully integrating people's perceptions into the assessment framework of local and national authorities.

People's perceptions, experiences, and issues in connection with particular places were collected in the two study areas – the Peak District National Park and the city of Sheffield. Tweets about places and stories of personal connection to places collected as short stories or longer interviews were collected and converted into documents for further analysis. The content and length varied within and across the data sources. Tweets were problematic as they tend to be short, due to an initial limit of 240 characters, and contain special characters, such as hashtags and user handles or emojis, to convey additional information in short form, which was not useable for the text-based analysis. Similarly,

---

[10] Heritage Futures: https://www.youtube.com/watch?v=oe0aMQXzqLI

[11] Everyday Heritage project, Australia: https://everydayheritage.au/

survey responses, where participants were asked to provide stories of up to 300 words, varied substantially in length and content. Some of the stories were one-word responses, while on the other end of the scale, stories could have a length of up to 303 words. The majority of stories had an average of around 100 to 150 words. More depth provided individual interviews with selected participants with a length between 4,351 and 12,773 words.

Users of the social media platform Twitter (now X) provided information on trending topics, issues, and favourite places in the National Park (**Chapter 2**). The analysis of almost 2,000 tweets across three bank holiday weekends gave insights into the emotional connection between people and the National Park. This provided an overview of topics but was limited in its potential to inform in more depth on social values and reasons for place attachment. Interview data provided deep insights into reasons for place attachment. However, the method did not allow a statement on wider landscape perceptions.

The most useful data source for this kind of research question and analysis proved to be the survey approach. The length of the stories in survey responses provided sufficient information to explore the reasons behind place attachment. Furthermore, the opportunity to target groups of residents within a bounded system or study area made the survey method the most effective data source for this research. The survey was able to capture narratives of connection to places in a bottom-up approach without a predefinition of what constitutes heritage and without limitations to places people were able to select as their most valued. This constitutes a distinctive difference from surveys typically conducted by Historic England (English Heritage 2000, 2) and the National Trust (2017), which usually limit the scope of opportunities for a people-centred assessment to predefined designated or otherwise officially recognised heritage assets.

The advantage of allowing participants the widest possible freedom in choosing valued places lay in the opportunity to enhance the understanding of the broad range of elements in the historic landscape, including natural and cultural features. Without restricting people's scope of replies, the gathered background information can help gauge people's appreciation of the heritage assets as defined by experts, as well as reveal elements of the landscape that matter most to people and have not yet been recognised by professional heritage managers or planners.

8.3.3 Social value versus individually held value

Social values, as defined by Jones (Jones 2017, 22) is the 'collective attachment to place that embodies meanings and values that are important to a community or communities'. According to the Burra Charter (2013, Article 1, Paragraph 1.2), stakeholders are identified as 'individuals or groups' who value places. There is increasing awareness of the significance of community and group inclusion in people and social value-based research projects, which were increasingly shaped by agendas and charters over the past 30 years (Council of Europe 2000; English Heritage 2008; Johnston 1992; Jones and Leech 2015; Jones 2017). However, the individual, as part of the whole, has so far not been recognised in the assessment of the historic environment or place-based decision-making of local authorities (Johnston 2023, 247; see also **Chapter 6**).

Several projects presented in **Chapter 1.2** showed previous attempts to integrate social values or people's perceptions in place-making and development decision-making (Primdahl and Kristensen 2016; Dalglish and Leslie 2016; see also **Chapter 5**), but the artistic expression of laypeople's opinions was deemed too challenging to be meaningfully used and integrated into the assessment frameworks (Dalglish and Leslie 2016, 217). Heritage practitioners continued to focus on expert expressions and assessment as defined by the Authorised Heritage Discourse (AHD, Smith 2006; also **Chapter 5**). Furthermore, as noted by Johnston (2023), the focus of understanding people's values in landscapes was very much on communities and groups, not the individual. The issue and challenges in working with communities or groups were emphasised by Dalglish (2018, 56-58). He argued that results of group assessments have a high potential of emphasising dominant voices or opinions within the group over marginalised, less dominant voices and opinions. However, Modesto and Waterton (2020) demonstrated (see also **Chapter 4**) the important role of private or individual opinions in heritage management and the importance of understanding personal connections.

As demonstrated in **Chapter 1.2** and **1.3** academic research, community-led initiatives, and organisation's engagement with the public, e.g., national parks, Natural England or Historic England, developed a wide range of tools and methods to recognise and realise the integration of social values in heritage and landscape management. However, none of these approaches focussed on the assessment of individually held values and their representation as 'hidden' value communities that has the potential to give underrepresented opinions a voice. Such background information can feed into currently existing projects, e.g., the Deep Cities or Social Value Toolkit (see **Chapter 1.2.7**) to identify both participants and places for in-depth social value assessment.

The exploration of individually held values in this research showed that people's personal stories of the connection and attachment to place, but also the challenges and issues associated with the environment, can be meaningfully translated into maps to convey local knowledge to a wider range of stakeholders (see **Chapter 4,** Figure 4-8 and **Chapter 6 ,** Figure 6-2). The following section will show how AI techniques were adopted, adapted, and applied across the different data sources, before I detail the categorisation of social values based on the stories of people and places.

8.3.4 Methods applied across three data sources

Methods used and developed in this research include Natural Language Processing (NLP), Named Entity Recognition (NER), Sentiment Analysis (SA) and Topic Modelling (TM). NLP provided the tool to prepare the data from all three sources for further analysis (textminR package in R after Jones, Doane and Attbom 2021). NER, a component of NLP, was more problematic as 'off-the-shelf' applications would not yield a good result in identifying and extracting locations from unstructured text documents (tweets, stories, interview transcripts). Algorithms are commonly trained on a variety of data but were not sufficiently fine-grained and place-specific to find places in the PDNP. It was, therefore, necessary to produce a dataset by filtering location data for the PDNP and creating a gazetteer (an index of locations), which could be provided to the algorithm for the automatic extraction of locations from unstructured text (for details see **Chapter 3,** Figure 3-10). This dataset was also used for the survey analysis, where participants were not able to use the map function to locate the points of their favourite place on the embedded Google map in the questionnaire but instead provided the name of the location (see **Chapter 4**).

Similarly, SA, commonly used in social media research and applied in social media analysis, proved to be useful for identifying negativity in the stories of survey participants, which allowed the detection of issues across the National Park (see **Chapter 4** and **Chapter**

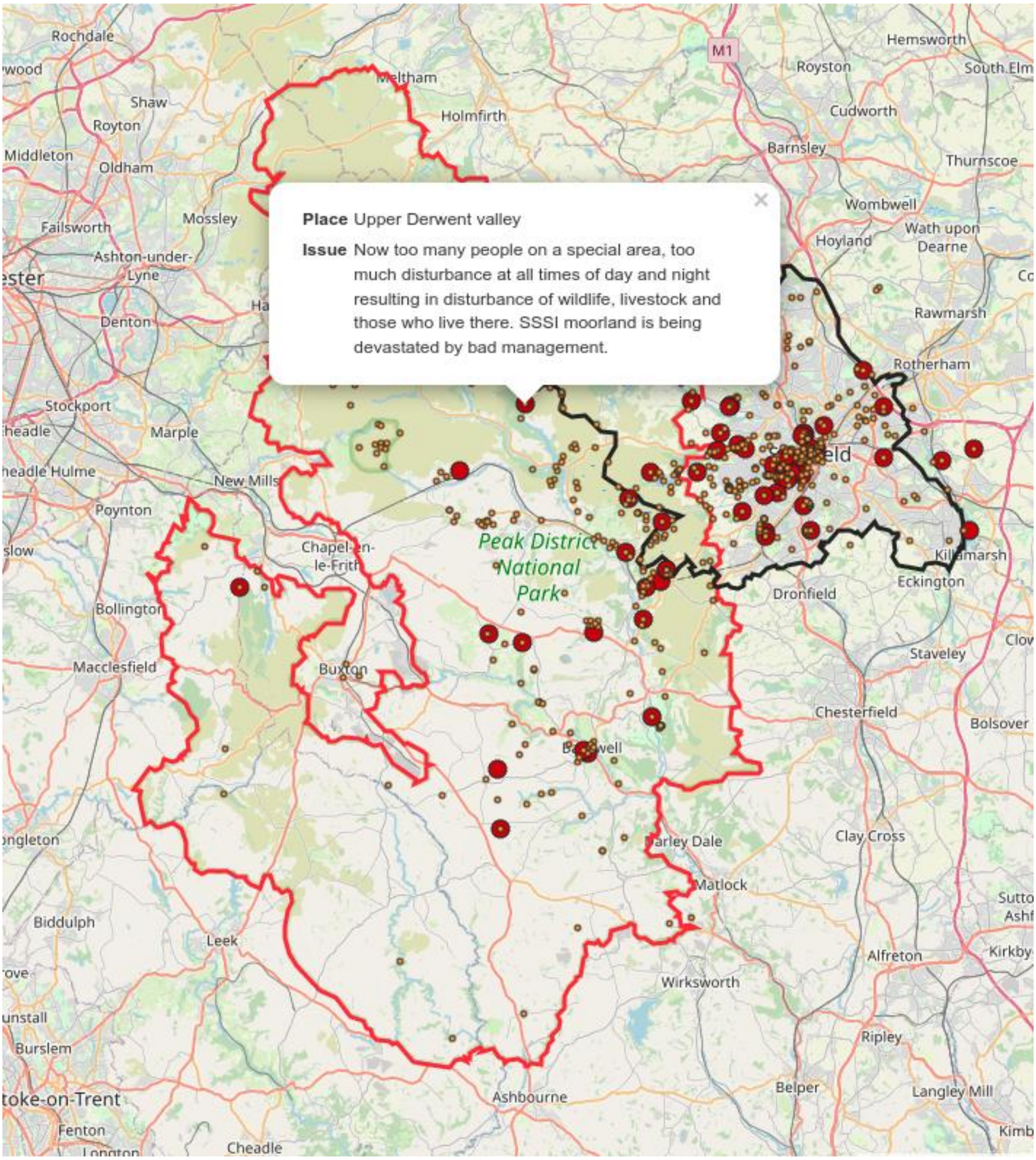

Figure 8-1: Extract of interactive webmap showing issues reported by resident of the PDNP and Sheffield associated with places in their stories. (Base map Openstreet map).

8, Figure 8-2). But as with many NER corpora, sentiment analysis developed based on US training datasets can introduce issues in text-based analysis outside the United States of America. The example provided in **Chapter 3** showed the difference between American

English and British English applied to the phrase 'quite good', which would be interpreted differently in emphasis in the US and Britain. Nevertheless, in general, the sentiment analysis illustrated a changing attitude of park residents and visitors before, during and after the COVID-19 pandemic both in the Twitter analysis and, subsequently, in the survey analysis. Tweets and survey participant stories showed clearly that this unprecedented event had huge social impacts on all aspects of life, including the connection between people and heritage and the natural and cultural landscapes, which was particular of interest for this research (**Chapter 3** and **4**).

Sentiment analysis based on text and emojis illustrated increased negative sentiment during the pandemic, which was visible in the frequency of negative emojis. For example, swearing and angry emojis were not represented in the years before and after the severe restrictions due to the pandemic, but were in 12th and 13th place in the frequency table during 2020 (see **Chapter 3**, Table 3-3). The emoji analysis also revealed a shortcoming in the text-based analysis due to the inability of the automated process to detect sarcasm in human communication. The text-based analysis looked much more positive in its result than the negative evidence from the emoji analysis (**Chapter 3**, see Figure 3-6) and the confusion matrix of False-Positives in the text-based sentiment analysis shown in Figure 3-7. SM applied to the survey data revealed issues and negativity, which were used to highlight developing problems in the study area as shown in Figure 8-2 (see also **Chapter 4**).

Topic Modelling was applied to the survey data and interview responses. Applying the method requires sufficient document content and length, which proved problematic with some of the one-word or short answers in the survey and made the tool not useful for the mainly short tweets[12]. The tool was, therefore, not used for social media text categorisation. However, TM proved to provide valuable insights into the topics of survey stories and interviews. The advantage of the method was that the empirical data was initially clustered and labelled thematically by the automated process, which sometimes revealed unexpected and surprising insights that were not initially anticipated. For example, an unanticipated strong emphasis on pro-environmental behaviour was identified during

[12] TM was not used in the social media analysis as topics or trends in tweets were inferred from hashtags which are commonly used to show trending topics.

the analysis of interviews **(Chapter 5** and **Chapter 8.3.5.4**). Similarly, the strong connection between local residents and the history of the landscape were anticipated, but the level of attachment to heritage assets was confirmed by the result of the automated process, unbiased by the subjectivity and anticipation of the researcher[13].

In summary, some methods used in this research for one data source proved also useful for another data source, for example, sentiment analysis for social media and survey data, but not for interview data. Also, Topic Modelling yielded good results for survey and interview data to aid the categorisation of themes latent in the data; however, it was not applicable to social media data due to the partly cryptic and short tweets. Emoji analysis was an essential element in social media analysis and might be a way forward to gauge sentiment in other data collection methods, for example, as shown in research using photo elicitation methods (see Madgin 2021). However, the method was not applied to survey or interview data. Developing a gazetteer for the NER process was useful for all three datasets for automatically extracting spatial information from unstructured texts and locating these on a map for an area-specific and fine-grained GIS analysis.

Furthermore, the online survey allowed for more targeted data collection focusing on residents and people working in the Peak District National Park and Sheffield. This focus enabled the categorisation of reasons for a connection and valuing places based on personal stories. Such a focus was not possible with social media[14] and added, therefore, to the breadth of information as background for the management of specific areas, such as, for example, national parks or cities.

---

[13] **Chapter 8.7: Summary of ethical implications, bias, and limitations** will detail the level of researcher bias and subjectivity introduced at various levels of the analysis. Some steps were more influenced by the researcher than other, for example the selection of interview partners, survey questions, and the manual assessment introduced researcher bias. The automated process of TM is based on statistical computation and therefore objective and replicable.

[14] Social media data was selected based on hashtag search focusing on the Peak District. However, due to the unreliable location information, it was not possible to select tweets from within or without the study areas (see **Chapter 3**).

Applying AI techniques to qualitative data has advanced in recent years (see **Chapters 1.5.4 and 5**). However, to avoid the 'black box' effect that can result from an NVivo approach, for example, from sentiment analysis or through 'autocoding', explainability of the algorithms and processes was paramount to ensure a good understanding of the performance and results of the AI-based analysis. The method provided offers a tool for social value analysis form unstructured texts that allow to find commonalities, issues and differences in the perceptions of wider landscapes.

8.3.5 Value categorisation from stories

8.3.5.1 <u>Critique of rigid value categorisation</u>

Categorisation as a tool to integrate social values into heritage and historic environment management has been critiqued and questioned in the past due to its dynamic and fluid nature (Jones 2017, 25; Rudolff 2006, 229; Smith 2006, 71, 80). National heritage organisations in Australia (see **Chapter 6**) and the UK (see **Chapter 4**) assess and manage the historic environment based on categories that include social values (Australian Heritage Council 2009; English Heritage 2008). However, as shown in **Chapter 2,** this category has not gained the same recognition in the official assessment strategies as did, for example, *Historical* or *Aesthetic Value*.

Extending the definition of the historic environment to the everyday and including laypeople in the process, shows that the successful application of principles of *Grounded Theory* in correlation with the *Conservation Principles* developed by English Heritage (now Historic England) can create dynamic categories developed based on empirical data (see **Chapter 4**). Categories developed based on the opinion of people who occupy the places, not defined by experts on the basis of historical, evidential, aesthetic and communal values (English Heritage 2008), provide essential insights to understand and improve places. Rudolf's (2006, 229) plea to see heritage from a people's perspective 'not as statements but as ongoing narrations' can be achieved effectively and efficiently using classical AI techniques as presented in this research. Similarly, Jones' (2017, 25) argument for a more people-centred approach that avoids the creation of categories as sameness and exclusion can be addressed by the SLC method developed in this research. Sameness arises from patterns across landscape based on the same experiences or perceptions of people, as 'hidden' or 'invisible' communities, which was previously identified as one of the key findings of this research (see **Chapters 4** and **6**). These people are otherwise not connected. The emphasis of this categorisation process, therefore, aims to identify value solely as a connecting element not as an exclusion tool.

The assessment and refinement of Historic England's *Conservation Principles* as a category system for assessing of the historic environment (see **Chapter 4**) will be detailed in the next section, followed by a free categorisation of the interview data and a comparison and correlation of the categories.

8.3.5.2 Categorisation based on an existing framework

Based on the survey data, I wanted to test to what degree the argument was valid that categories in a predefined assessment framework are not fit for purpose to accommodate the current, individually held values of laypeople – or insiders – as opposed to a fact-based, 'objective' assessment of heritage experts – or outsiders – who designed the framework. To assess if the stories, experiences and perceptions of people could be categorised based on an existing assessment framework, which is in practical application for more than 15 years, I applied the systematic approach of the *Conservation Principles* (English Heritage 2008) to the stories provided by survey participants. As shown in **Chapter 4**, the approach included the use of NLP and TM to reveal themes latent within the data. The emerging themes were labelled based on the most frequent big-grams extracted from the stories provided by the survey participants and, subsequently, correlated and adapted to the existing category framework (see **Chapter 4**, Figure 4-5, Table 4-2).

The analysis revealed that people's reasons for valuing specific places could be categorised using the existing system of the *Conservation Principles* up to a point. While *Historical* and *Communal Values* dominated based on people's stories, *Green Spaces* were identified as an additional category of value as an essential part of the environment in the perception of people. This was emerging from the data and not predetermined or anticipated at the outset (see **Chapter 4**, Table 4-2).

This holistic perception of people, as evident from their stories, shows that nature and culture should be seen not as separate parts of the environment but in a holistic view.

This constitutes a second key finding of this research and was elaborated on in the introduction (see **Chapter 2.2.1,** The 'old' model, and **Chapter 5**). This insight shows that the current discourse on the artificial dichotomy of nature and culture in heritage management is essential for the positive management of places from the viewpoint of laypeople (for the discussion see Byrne and Ween 2015; Harrison 2015). The dualism of nature and culture has been seen as a direct link to behaviour and morality in regard to nature and can have an impact on the future discussions about the role of heritage in tackling the pressing issues of the time, such as climate change (see e.g., Bryant and Farrell 2023; Scannell and Gifford 2010). *Green Spaces* as value category was further divided into the subcategory *Health*, as people valued their environments for the benefits for physical and mental health and wellbeing (**Chapter 4**,Figure 4-7).

The subcategories for *Communal Values* were expanded and refined, as this category, as defined by the *Conservation Principles,* was found to be not sufficiently fine-grained to accommodate people's individual stories and, therefore, encompass the various reasons behind individual values and place attachment (**Chapter 4**,Table 4-2). The original subcategory of *Social Values* was renamed *Private Values* in order to avoid confusion with the international definition of social values used in the Burra Charter (see Johnston 1992; Jones 2017, 22). The emphasis of this category lay solely on the individual and personal connection between a person and a place, which included a wide range of personal reasons, such as the personal family history. Furthermore, the subcategory *Arts & Culture Value* was added, since intangible cultural assets, such as theatres or museums, merit inclusion into the canon of place values (**Chapter 4**, Table 4-2, Figure 4-7). This notion feeds into the discussion about tangible and intangible cultural aspects of place – another dichotomy that was found to not exist in people's perceptions. Again, the holistic perception of the environment became apparent in the stories provided by survey participants, which comprised a wider range of values embedded in single places.

The survey provided essential information on people's reasons for a deep connection, issues and needs associated with places. Some of these reasons were personal life history and memories, for example, the church of their wedding, the workplace of a father or childhood memories of playing in the botanical gardens; the connections to the local history, for example, the industrial past in Sheffield with steel works and mills or the spiritual connection to places of ash scatters of loved ones memorising their love to places in the national park. The stories revealed a variety of reasons behind place attachment, which allowed and required a categorisation that provided the same freedom to accommodate the fluid and dynamic quality and the holistic view of people. This background will help communities make a case regarding planning decisions and allow local authorities and national organisation to act proactively to planned change and development. The following section will detail how the categorisation of interviews evolved from within the data without a predefined categorisation system.

8.3.5.3 Categorisation based on Topic Modelling labels

As shown above, matching people's values ascribed to natural and cultural landscapes works to some degree by adding additional categories to accommodate specific elements not recognised or covered by the existing framework of the *Conservation Principles*. However, to address the argument that social values or individual values are too dynamic and fluid and should not be fitted into a rigid existing framework, a narrative-based approach tested the categorisation without a predefined category system on the semi-structured interviews (see **Chapter 5**) in order to, subsequently, correlate the categories of both approaches for comparison.

The interview data was analysed using the same method as for the survey stories (see workflow of **Chapter 4**, Figure 4-1). The longer interview transcripts were divided into smaller text blocks, pre-processed in NLP and analysed using TM. The themes emerging from the data were, subsequently, manually summarised, ordered in categories and subcategories based on the labels provided by the algorithm and presented as a category framework for this specific analysis (see **Chapter 5,** Figure 5-1).

The emerging categories were developed in a narrative-based approach that encompasses all aspects of the reasons for their sense of place and belonging. The category labels were based on the topic labels that were identified during the TM analysis (the themes and labels with keywords of the clusters are presented in **Appendix for Chapter 5/Supplementary Material 2**, the correlation of topics generated by the TM process and the manual observation can be found in **Supplementary Material 3**).

The results of this narrative-based categorisation approach are presented Figure 5-4 (**Chapter 5**) and on the left-hand side of the correlation graph in Figure 8-3 and show that *Landscape Quality* and *Change and Continuation* were the predominant themes in people's perception of places in the living and working landscape of the PDNP. This is followed in third place by the surprising insight that environmental issues and proactive initiatives and behaviour rank high on people's perceptions. For example, in Interview 2, the notion of 'reinstating former hedges' and 'reintroducing biodiverse hay meadows and old apple species' was an important aspect of landscape perception. Furthermore, Interview 7 mentioned the initiative of 'renaturalisation' and 'less use of weed spray' to enhance the natural environment.

Further topics emerging from the analysis included the connection to the local history or personal history, with the category *Place History*, and strong community and people-centred topic summarised in the categories *People and Place Engagement* and *Communities*. However, the perception of landscape is not always positive, a notion also found in the analysis of social media and survey data. Categories reflecting the issues and challenges associated with places in the PDNP were thematised in the category *Challenges*.

In summary, the case studies provided a flexible categorisation of stories that supported the understanding of the place attachment of residents living and working in the national park. The analysis following the principles of *Grounded Theory* allowed the themes to emerge from the stories. It revealed topics, such as *Pro-environmental Behaviour* and *Challenges*, which were not anticipated *a priori* and are not included in the assessment framework. The following section will correlate the findings of both approaches: the predefined categorisation system based on extended categories of the *Conservation Principles* and the narrative-based approach.

8.3.5.4 Correlation of pre-defined and narrative-based category systems

The correlation of value categories in Figure 8-3 illustrates the differences and the similarities where topics in both the survey and interviews matched in some categories. The first impression of the graphic illustrates that the interpretation of categories from the interview data, with an interest in the varying reasons behind a sense of place, is less rigid and strict, allowing for more flexibility. The left-hand side represents the categories developed based on survey responses in conjunction with Historic England's *Conservation Principles* (English Heritage 2008) and extended categories, as detailed in **Chapter 4**.

The right-hand side represents the synthesised categories that emerged as themes or topics from the interview data (see **Chapter 5**). The categorisation of the survey data matched Historic England's value categories for the assessment of the value of a place. Historic England[15] stresses the importance of a systematic and consistent approach to value assessment, which can aid the decision-making process in the historic environment. At the same time, researchers, for example, Rudolff (2006, 229) argue for a narrative-based dynamic system, similar to Stephenson's 'relationships, forms and practices' (Stephenson 2008, 134 -135, and 134, Fig. 2). Stephenson recognised the necessity to add natural, community, and social values as well as intangible aspects of the value process, such as traditions. However, a time and resource-sufficient technique had to be developed to achieve such an assessment. AI tools, such as Topic Modelling, allowed the narrative-based, individual value assessment as proposed in this research, which puts the values conveyed in the personal stories of the people at the heart of the assessment and not the predefined categories.

The correlation of the two resulting categorisation systems reflects the holistic perception of the landscape that occurs in people's minds when developing a sense of place and attachment. This insight made the additional value categories – added to the *Conservation Principles* – like Stephenson's findings, necessary and mirrored the more complex category system of values emerging from the interview analysis. The interview analysis aimed to gain a deep understanding of the reasons behind place attachment and sense of place. Therefore, the main and subcategories were more varied and fine-grained. The main categories of both systems show a high congruence in, for example, *Historical Value* and *Place History*, which reflects the association of people with the rich historical past of places. The interviews and survey participants either had a personal connection to historic events, objects and sites or felt a deep appreciation and connection to the history of places. For example, such connection was represented by the drystone waller Trevor (**Chapter 5,** Interview 1) with a deep connection to the drystone walling tradition in the Peak

[15] https://historicengland.org.uk/advice/constructive-conservation/conservation-principles/

District and an appreciation of the skills in honour of the achievements of past generations. Similar examples were found in other interviews and survey responses (see **Chapters 4** and **5**) and reinforced the notion that this category exists in both, the internal, individual value system of people and in the fact-based expert-led assessment system, simultaneously.

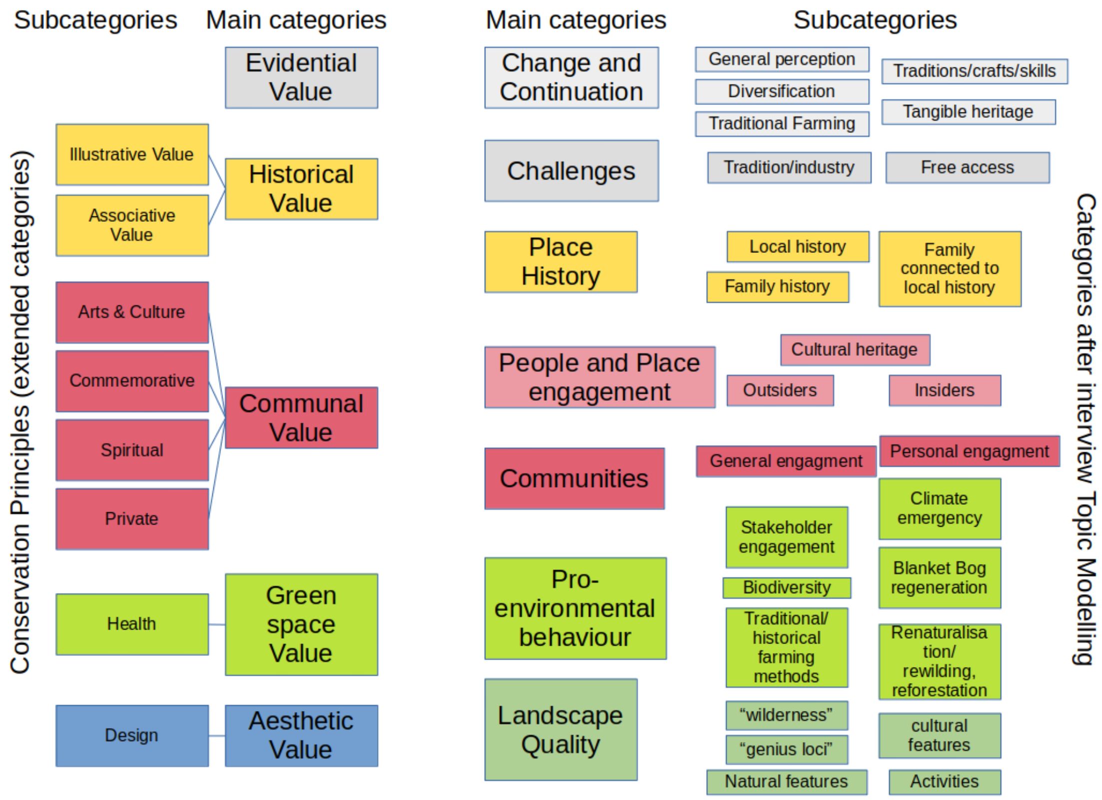

Figure 8-2: Correlation of categories as defined after the Conservation Principles (English Heritage 2008), including additional categories based on the survey data analysis (Chapter 3)(left-hand side), with categories identified in the analysis of interview data (Chapter 4) (right-hand side). Individually held values categorised based on a system (semi-predefined) and free categorisation not based on a predefined system show similarities and overlap.

Similarly, *Communal Values*, including *Private*, *Spiritual*, *Commemorative* and *Arts & Culture* elements, were matched in the themes emerging from the interview data as categories of *Communities* and *People and Place Engagement*. These communal or community-focused elements of valuing place are less important in the expert-led valuation system, as shown in **Chapter 2**. Both survey and interview data have shown that these values matter to people. In reaction to this result, I proposed an extension and refinement of the *Communal Value* category in Historic England's *Conservation Principles* (**see Chapter 4**).

A strong emphasis on the category *Landscape Quality* emerged from the interview data, and with the subcategories focused on the aspects of wilderness, genius loci, natural and cultural features, and activities, this can be matched with the equivalent of both *Aesthetic* and *Green Space Value* in the extended *Conservation Principles* system. All of these categories include views, the beauty of natural and cultural features and overlap with the category of *Pro-environmental Behaviour*, with its focus on the natural aspects of the landscape, with elements of concern for biodiversity loss and climate change, as well as blanket bog regeneration, rewilding, renaturalisation and reforestation.

Less dominant was the awareness of the category of *Evidential Value* in the *Conservation Principles* system among people's perceptions, probably less relevant due to a lack of awareness or opportunities for use and relatability. Similarly, on the side of the interview-data-generated category system, the themes of *Challenges*, reflecting issues in current heritage and landscape management, and *Change and Continuation*, illustrating a changing world and the need to preserve some elements of natural and cultural landscapes, had no equivalent in the survey categorisation system (Figure 8-3). The latter categories reflect the negative sides associated with heritage and landscape management, which has no place in the established assessment framework. However, it is, for example, recognised through Historic England's *Heritage At Risk* programme[16]. Nevertheless, the ability of the categorisation approach based on people's perception has the potential to encompass the negative and positive sides of heritage and landscapes.

In summary, the correlation of categories inferred from survey and interview data provided a basis for creating a systematic framework for assessing heritage and the historic environment, including nationally and locally designated, as well as unrecognised everyday heritage, which constitutes the living and working environment that people value. Reasons for these values can be categorised in a narrative-based approach, using Topic Modelling to capture the latent themes and topics within the data that provide the overarching categories resulting from this analysis. Depending on the depth of information and detail, both the adapted *Conservation Principles,* as well as the non-predefined categorisation provide a useful approach to systematisation. Both systems showed similarities in

---

[16] https://historicengland.org.uk/advice/heritage-at-risk/

overarching categories and themes, with a more fine-grained subcategorization in the free categorisation approach. Based on these results, the categorisation and correlation successfully provided a method for systematisation of categories from narratives based on the principles of *Grounded Theory* and developing themes from within empirical data[17].

8.4 Objective 2

8.4.1 Assessment of HLC principles for SLC application

Historic Landscape Characterisation (HLC) provided the starting point for the approach in this thesis as a map-based support for change and development decisions in wider cultural landscapes. The intention was to adopt and adapt the key principles of HLC and integrate the individual values of laypeople as a counterweight to the expert-led and fact-based dataset HLC represents. The key principles of HLC were: present-day landscapes, not past; general, not particular; value-free; continuous coverage, not sites or point data; dynamic and updatable; understanding landscape through people's perceptions rather than as an objective thing; use of present-day landscape as source; reading landscape as material culture; semi-natural and living features are part of the landscape as well, all aspects of landscape matter not just 'special' areas (Aldred and Fairclough 2003, 31, 40-41).

To assess if the objective was achieved and if this was possible and realisable, it should be established first what the HLC method itself achieved and failed to provide based on the key principles compared to the actual outcome. The key principle that was achieved was, for example, the generalisation of landscape character. Even though different projects of HLC varied and were not consistent and compatible with each other, great effort was made to create a consistent national project in 2017 (Exegesis and Locus Consulting 2017). However, the national dataset had to compromise on the level of detail for the sake of applicability. At the same time, HLC aspired to express the distinctiveness of landscapes to enable adequate management of landscapes based on the elements that shaped the historic environment in every place (Fairclough 2007, 84). The aim to achieve distinctiveness

---

[17] The detailed lab book leads through the process developed as a workflow for analysing survey responses (see **Chapter 4**) and can be found in the **D. Appendices for Chapter 8: Appendix 7**.

contradicts the aspiration of generalisation. Therefore, SLC, developed in this project, focused on expressing distinctiveness rather than generalisation, despite the approach to achieve a categorisation of social values (see above and **Chapters 4** and **5**).

Another key principle achieved within HLC was capturing past processes in the present-day landscape. The focus of HLC is the landscape, as it can be seen today. Also, the landscape was read as material culture in the form of processes manifested in landforms visible as, for example, fields, hedges, woodland, and settlements. However, this contradicted the notion of not seeing the landscape as 'an objective thing'. HLC treated the landscape objectively only to a degree of interpretation of the experts who worked on the HLC projects, compared to subjective perception.

The strong side of HLC was the representation of the landscape as an area consisting of repeating landscape types. This concept fulfilled the principle seeing the landscape as continuous, not as a collection of points and sites. Furthermore, HLC integrated semi-natural landscape features. However, as shown in the introduction, some shortcomings of HLC were, for example, a bias towards particular landforms (see critique and discussion about HLC bias towards deer parks in contrast to fox coverts in **Chapter 2.6.2** (Finch 2007, 377)). Also, HLC achieved the aim of generalisation in their typology and value-free representation – not valuing one part of the landscape above another and including all aspects of the landscape not just the 'special'.

Were HLC failed to deliver, was in the creation of a dynamic dataset that could be updated and adapted. As Turner (2007, 46) pointed out, heritage management needs to adapt to the dynamic nature of landscapes. However, HLC as a tool was never updated and, therefore, represents a snapshot in time, contradicting the notion of a dynamic tool. No overarching methodology was created that allowed time- and resource-efficient data collection on a rolling basis.

The method also failed to present a landscape based on people's perceptions as initially intended (see **Chapter 2.6.2**; see also Clark, Darlington and Fairclough 2004; Dalglish and Leslie 2016; Turner 2018). The reason for this was most likely that this contradicts the other main key principles of the method, for example, value-free data representation and using complex jargon, making it valuable as a landscape history characterisation tool, but not for the integration of social values. Similarly, the integration of social values based on these principles posed unsurmountable problems for this research, which aimed to include people's perceptions in the existing HLC approach.

Based on this analysis, the integration of social values or individual perceptions of laypeople into HLC is not reconcilable with other key principles of the HLC method. Therefore, the principles and methodology of SLC need to divert in parts from the HLC methodology. Hence, a new methodology, building on the HLC, and a map-based tool were developed, which could be used in conjunction with HLC. For example, continuous coverage is not feasible in a social value approach. While blank areas in the SLC dataset do not necessarily represent areas with no value, but rather areas that lack data, continuous coverage is not the aim of the method. The method provides a background on places that matter most to people and full coverage is, therefore, neither feasible nor desirable. As the method should be applied on a rolling basis to collect more data, the coverage would increase, and the blank areas might shrink or shift over time. At the same time, surveys conducted on a rolling basis would also achieve the key principle of HLC – being dynamic and updateable.

SLC can also never be value-free. Social values are associated with places or parts of landscapes that have a deeper meaning for people than other places. The HLC principle contradicted the notion of integrating people's perceptions, which has proven to be highly value-based. People give meaning and value to places and landscape, which cannot be value-free. Similarly, the approach to identify places that matter most to people was conducted on a point-based approach, with the consequence of the return to point or site-based data. People can appreciate wider landscapes, especially views and routes for walking or other exercise. Nevertheless, relating to place will always be focused on specific parts of the landscape or sites, buildings, objects, or places associated with intangible heritage.

The correlation between HLC and SLC principles, as described above, would benefit from more data and analysis that lies beyond the scope of this thesis. The correlation itself proves problematic due to the different typology and terminology, which was also identified as an obstacle for integrating people's perceptions into HLC by Clark, Darlington and Fairclough (2004, 6) and Dalglish and Leslie (2016, 215). Due to the different approaches and intentions of HLC and SLC, it would be interesting to see a future exploration of the question of whether a correlation would provide evidence for causation or, in other words, if we could infer a specific category of social values from a specific characteristic historic landscape type identified in HLC. A correlation between HLC and SLC showed that specific landforms might influence present day appreciation of the landscape (Figure 8-4). However, more research is needed to strengthen the evidence for such an assumption. The next section will show this correlation between HLC data of the PDNP and the data of the three sources for the creation of the SLC dataset.

8.4.2 Correlation of HLC data and SLC hotspots

HLC presents the processes and practices in the past that led to the distinctive character of landscapes in the present in map format. SLC represents the spatial distribution of people's perceptions in today's landscape. This raises the question of whether the correlation between HLC and SLC can provide insights into the development of place attachment based on the distinctive landform types identified in HLC. Figure 8-4 shows the correlation between HLC data as polygons and SLC data as points.

The correlation shows landscape types that are more associated with social values identified in social media data (**Chapter 3**), survey data (**Chapter 4**) and interview data (**Chapter 5**). The visualisation of the area of Sheffield City has been excluded for this correlation, as the typology in urban areas is commonly extremely fine-grained and detailed. This area would make statements about the correlation of types in HLC and categories in SLC challenging and would be a field for future research. The larger polygons provided in rural areas, such as the Peak District National Park, allow a more general assessment of the correlation. The map shows that the northern part of the Peak District, also known as Dark Peak, has no notable social value attached to places in the area[18] and is seemingly less attractive to afford connection or attachment (Figure 8-4, *Open waste and commons*). The area consists primarily of blanket bog or moorland with no architectural features, such as drystone walls, field barns, or other landmarks and no notable above-ground-raising vegetation, such as small bushes or trees.

[18] It is important to stress that blank areas illustrated in this research represent areas with missing data not necessarily a lack of social values. Further research and annual reviews could improve the data basis and enhance the knowledge on social values in these areas.

Daniel Defoe, writing a travel diary on his journeys across Britain between 1724 and 1727 (University of Portsmouth 2017, online), describes the moorlands of the Dark Peak, famous for its dark grey gritstone geology, as 'perhaps the most desolate, wild and abandoned country of all of England'. A similar notion seems to make this landscape less attractive for bonding or attachment and less attractive for visitors and tourism today. However, the value of this landscape was elaborated on in the interview discussion in **Chapter 5**, where blanket bog regeneration featured as an important element for biodiversity. In the social media and survey data, there is, however, a notable accumulation of social value along the edges of the moorland, which is easier to reach and better connected to parking and roads (the image in **Chapter 5**, Figure 5-7, gives an impression of the landscape with the exception of some rare features, such as the flagstone path and the trig point). A typical hotspot in the west of the PDNP shows Kinder Scout, which was noted on the map in **Chapter 6**, Figure 6-2, and represents a landscape with favoured views and rock formations.

Area types, such as *Recreational, Reservoir* or *Urban*, show a high level of attraction visible as hotspots on the correlation map (Figure 8-3; see also in **Chapter 6,** Figure 6-2).

Notably, the areas with the landform type *Ancient Enclosure – Fossilised strip System, Medieval Fields* and *Irregular Fields* show higher attractivity for social values. The reason, also shown on the map in **Chapter 6**, Figure 6-2, lies in the proximity of such fields to medieval settlement cores, which have developed into the typical 'chocolate box villages' and market towns, with high attractivity for visitors and locals alike. Examples in the national park are Bakewell and Castleton (see image in **Chapter 5**, Figure 5-6).

The wide landscape of *Post-medieval Enclosures* and *Enclosed Moorland* in the south of the Peak District, also known as White Peak, mirrors the blank areas similar to the moorland in the north. This area is dominated by large swathes of upland pasture, again with fewer amenities and landmarks and, therefore, potentially less attractive to both visitors and residents of the national park.

To conclude, the correlation of HLC and SLC presents interesting insights into the potential impact of former land use and the processes that created the present-day landscape. Higher attractivity of landscapes around medieval settlements and today's market towns and Peak District villages is notable. This may have to do with cultural offers, amenities, and opportunities for community activities – reasons that were given in the interviews, survey responses and which were apparent from tweets. While moorland and other types of landforms recorded in the HLC dataset are important for biodiversity, which was thematised in in-depth interviews (**Chapter 5**, Interviews 9 and 10), or river valleys important in the survey data for recreation and health reasons (predominantly mentioned as reasons for value in the survey responses, see **Chapter 4**), such statements cannot be

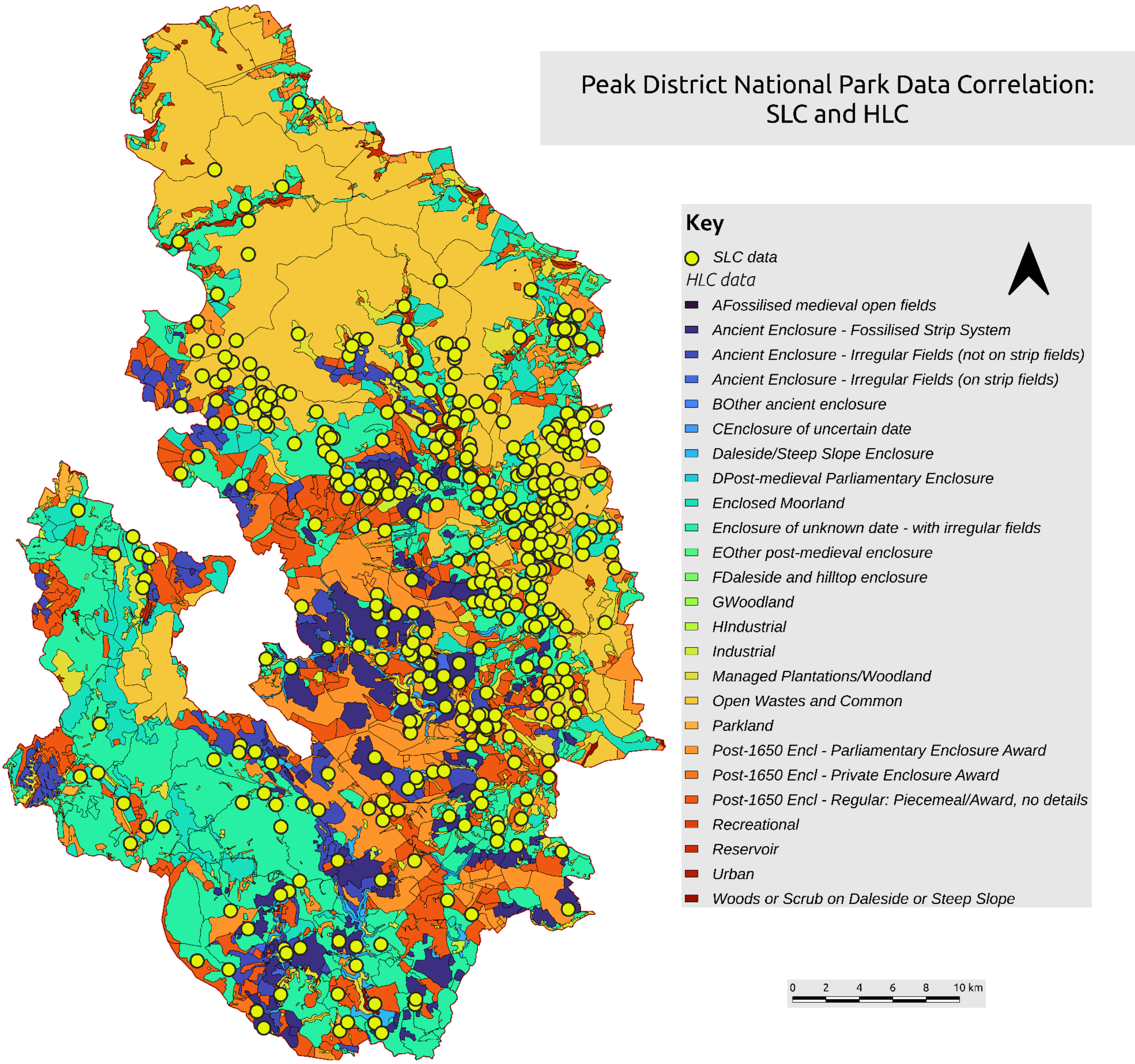

Figure 8-3: Correlation of HLC and SLC datasets, showing the land use typology, identified during the HLC categorisation of the PDNP. The correlation can show potential influences of practices and uses of the landscape in the past on present-day appreciation and the distribution of social values mapped from the SLC dataset (map contains HLC data provided by ADS).

inferred from HLC maps. Therefore, a correlation and meaningful comparison would need more research and more data to work from.

Nevertheless, the notable influence of medieval features, reservoirs and river valleys on present-day social values provides essential information as background for managing such landscapes. For example, when correlated further with the issue maps created from the SLC dataset, this information can predict areas that might have similar issues but lack data. Such areas could, in future, be targeted with surveys to assess challenges and opportunities for the management. For example, the issue maps show challenges in some areas, identified based on the stories of survey participants, as, for example, increased footfall of visitors, parking issues and undermining community coherence (see **Chapter 6; Chapter 8**, Figure 8-2). Further research can aid a broader understanding of the impact of past land use practices and current social values.

8.5 Objective 3

8.5.1 Patterns of 'hidden' value communities

The main focus of Objective 3 was on the output and visualisation of individually held values represented on maps, which could be used for public engagement, information and as background for planning and decision-making.

As seen, individual opinions are subjective and varied but provide essential insights into public perception, which group or community-based data collection cannot provide. Collecting individual opinions, stories and perceptions can be the starting point to finding commonalities and, subsequently, attempting the categorisation of individual stories. This may also form the basis for more site focussed projects on social values and identifying interest groups with similar landscape and place preferences.

A key outcome of this thesis constitutes the development of a methodology that allows collecting individually held values and categorising these. The final step of this analysis is the visual representation of these values or categories across wider landscapes. GIS provides a flexible and versatile tool to convey such information effectively. The maps presented throughout the publications included in this thesis show that individually held values can form patterns across wider landscapes**. Chapters 3** and **4** illustrate such patterns, which represent ‘hidden’ or ‘invisible’ value communities (also **Chapter 6**). The analysis of the stories presented a method to allow individual stories and experiences to be mapped and to form natural communities or categories based on the same place (place-based values) or same reasons for connection to place (story-based value).

The emerging patterns visualise the special places that matter most to people and include grand and designated heritage, as well as the mundane and everyday elements of the environment that are not officially designated as heritage. In the past, an attempt to include people’s perceptions in heritage assessments was deemed too challenging, with a risk of the professional field losing the domain on the management and assessment of heritage (see **Chapter 2**; see also Dalglish and Leslie 2016, 217; Jones 2017, 28; Smith 2006). However, emerging patterns provide a tool for communication between heritage professionals and laypeople. Communication and meaningful integration of local knowledge and expertise can help build a resource for professionals and practitioners. This background information provided in maps and databases showed hotspots (**Chapter 3,** Figure 3-12**,** Figure 3-13 and Figure 3-14), categorisation (**Chapter 4,**Figure 4-8), stories (**Chapter 4**, Figure 4-3) issues (**Chapter 5**; **Chapter 8**, Figure 8-2), sentiments of the public towards places (**Chapter 3**,Figure 3-5) and opportunities for public engagement (**Chapter 8**, Figure 8-5 and Figure 8-6). Against the background of previous projects of artistic expression of distinctiveness and character as shown by, for example, Common Ground (Common Ground 2006, 1996) and initiatives to incorporate people’s perception in local decision-making, this research has developed a methodology to collect and analyse individually held values and represent these as patterns across wider landscapes through SLC. The category- and place-based SLC method provides an opportunity to meaningfully assess and integrate people’s perceptions into official assessment frameworks for informed and proactive heritage and landscape management.

8.5.2 Correlation of favourite places

Both survey and social media analysis provided point or location data, which could be visualised and analysed in GIS (see **Chapter 6**, Figure 6-2). While the data sources were treated as separate datasets, a correlation showed that, for example, the distinctive pattern of places mentioned in tweets and favourite places mentioned by survey participants differed in some areas, while some localised hotspots matched. **Chapter 6** presents the findings of the three data sources (**Chapter 6**, Figure 6-2) to allow the comparison of locations mentioned by Twitter users from across Britain[19] and survey participants located in the PDNP or Sheffield. Notable are the areas where locations were appreciated for their qualities in both tweets and stories provided by the survey. These places were, for example, Kinder Scout and Stanage Edge, moorland plateaus that afford wide views and opportunities for outdoor activities, Rivelin Valley (**Chapter 6**, Figure 6-5), and Padley Gorge (**Chapter 6**, Figure 6-4), a rare temperate rainforest, both with natural qualities and opportunities for recreation. Typical Peak District market towns and villages offering a wide range of amenities and history were similarly favoured by both study groups and included Bakewell and Castleton (**Chapter 5**,Figure 5-6). Furthermore, Chatsworth House (**Chapter 6**, Figure 6-6), a stately home with extensive parkland, and Redmires Reservoir, one of the artificial water bodies in the national park, attract people and invite them to partake in outdoor activities and recreation.

In contrast, the two data sources also show distinctive differences in the distribution of location in relation to the wider landscape. Hotspots, based on tweets and survey responses, did not coincide in some parts of the Peak District. For example, the map illustrates that places mentioned in tweets were more likely located in the White Peak to the south, while survey respondents connected more to places in the Dark Peak to the north and closer to Sheffield[20].

---

[19] While the location information provided by users themselves does not give reliable information on the location of Twitter users, there can be only an approximate location of the tweeters (see **Chapter 3**).

[20] It should be noted that the social media research was focused on the PDNP area, to limit the data to a manually assessable dataset, while the survey was conducted in the PDNP and the city of Sheffield. The responses of the survey were, subsequently, filtered to the PDNPA area but included responses from residents of both study areas. This could be a factor for the density of points closer to the Sheffield city boundary and might be an artefact of the small dataset size with a potential to converge with an increased data base. However, overall, the data allows to identify hotspots and compare the two data sources.

Interview participants were spread fairly evenly across the study area to provide insights into various areas from the perspective of an ‘insider’ or a person with an individual and particular connection to the place. This detailed information provided a deeper understanding of the area, potential reasons for place attachment and the distinctive quality of place (see **Chapter 5;** for distribution of interviews in relation to other data sources see **Chapter 6**,Figure 6-2). Such reasons were, for example, the local history as a connection between retired farmer and cultural centre manager Elspeth and Pilsbury castle, a motte and bailey castle (**Chapter 5**, Interview 2). As another example, Sue, a retired farmer and artist has a particular connection to the history of her farm and the landscape surrounding it, particularly a part of her farm that contains a Roman-British settlement (**Chapter 5**,Interview 7). Also, the personal history narrated by Joan as one of the descendants of the survivors of the bubonic plague in 1665 (**Chapter 5**, Interview 5) or the deep connection that the distinctive landscape features such as drystone walls (Trevor, drystone waller, **Chapter 5**, Interview 1) played in place attachment. This additional information helped to gain a deeper understanding of the hotspots and the landscape as a whole.

The correlation between the data sources provided a broad overview and a deep understanding of the locations and reasons behind the attraction to specific landscapes, favourite places, issues, and challenges. This shows the variation of places that matter to people in the Peak District and their meaning based on individual perception and experiences. At the same time, blank areas show places where no data exists, which raises the question of whether these blank areas are actually deprived of value and lack an affordance for meaning making. What are the opportunities in these areas, and how can challenges in the highly favoured places be tackled? This background information can enable proactive heritage and landscape management for informed decisions based on the stories of residents and visitors to the area.

8.5.3 Integrating individually held values in official assessment frameworks

The publication in **Chapter 6** provides the background of how the objective was achieved to include social values into official assessment frameworks by creating a database and a series of outputs for the visualisation. Using GIS to create a range of maps serving different stakeholders allows a convenient medium to connect to similar datasets, as shown in the example of HLC. **Chapter 6**, Figure 6-7 illustrates a map stack as a framework of maps that differs in their potential for engagement with the public or complexity to aid the work of local authorities. This framework can be integrated into the GIS of local authorities and correlated or used in addition to existing maps for landscape-scale or site-based decisions. Existing map-based assessment frameworks, like HLC, which have been used to aid such decisions on change and development in natural and cultural landscapes, have a high complexity and provide information on past processes and land use that shaped the present-day landscape. Such datasets are valuable for informing the planning process on how much change a landscape could absorb without losing its typical character. However, as shown, HLC maps do not include people's perceptions and what they deem valuable as a distinctive feature of the landscape.

To complement the factual, historic basis of information on landscape, this thesis proposed the SLC approach to include the aspect of individually held perceptions, values, and meanings of laypeople in landscape assessments. This combination of facts on historical land use and the people-centred characterisation approach has the advantage that change and development can be assessed and managed against the background of local knowledge and public acceptance. As an example, specific to national parks, local park authorities are better equipped to identify areas highly valued by residents as opposed to areas that afford less potential for a personal attachment when deciding locations for new quarries (see **Chapter 1.2.3**). In general, local authorities can use SLC to gauge the local sentiment towards, for example, residential development or infrastructure projects. For instance, a long-running dispute was the selection of the Old Oswestry Hillfort site for a residential development[21]. Local authorities underestimated the value and the resistance of the local population. SLC information provides a background that enables local authorities to assess planning decisions proactively. **Chapter 4**, Figure 4-8 shows an example of how the survey data can be used to identify areas with the highest degree of value associated with places and what the reasons are for favouring these places (category map).

Chapter 6, Figure 6-7 illustrates a framework of hotspot, issue, and category maps, which is based on the SLC data collection and analysis method developed in this research.

---

[21] http://oldoswestryhillfort.co.uk/press-release/day-of-reckoning-for-shropshire-hillfort/

SLC offers opportunities to query the dataset and produce outputs that can address a range of questions about the landscape for informed heritage and landscape management.
For instance, an important aspect that GIS maps can support is the visualisation of issues and negative experiences of people in relation to the environment or identify places with high value, where the acceptance of change would be lower[22]. The proposed SLC tool provides local authorities with opportunities to plan and manage heritage and cultural landscapes socially sustainably [23]. However, the database also offers opportunities for public engagement. The following section will extend this opportunity to provide inspirational maps as products for active public engagement and benefit.

8.5.4 Output for public engagement

Chapter 3 provided visualisations that are less complex but more engaging, for example, the emoji maps visualising the sentiment of Twitter users towards places in the Peak District (**Chapter 3**, Figure 3-5). Maps for public engagement should convey less complex information and use less abstract visualisations. The story map (**Chapter 4,** Figure 4-3) presents the stories of survey participants as a way to let the public or outsiders partake in the experiences of local people and immerse themselves in the landscape based on the stories and photographs provided by the survey participants that shared their favourite places.

As a further example of such public engagement opportunities, self-guided walks were created as part of this research and presented during an outreach project[24]. The routes of the self-guided walks connected locations provided as favourite places by the

[22] The question of the degree of accepting change in their favourite places was part of the survey and can also be queried in the SLC map.

[23] The proposed SLC tool for integrating social values into the assessment framework was developed for the PDNPA and presented to the local authority. The methodology will be provided to the local authority for application in a real-world management environment. The testing of the method was beyond the scope of this thesis due to time constraints but is anticipated by the PDNPA in due course.

[24] The project website can be found at: https://peopleandplaces.uk/

survey participants. They included the personal stories and comments of the participants, allowing the interested public to explore the places to which local people have connection or attachment. This also provided some interesting insider information that is not presented in official guidebooks. This very personal exploration of the landscape is presented as leaflets (Figure 8-5 and Figure 8-6) and was also available as walking maps.

The opportunities to use the data set of individually held values allows to communicate place attachment to the public and create a basis for cooperation and engagement.

8.6 Summary of the discussion

The objectives of this research were achieved using and developing innovative methods with a focus on the individually held values that people place on natural and cultural landscapes. The project started from the HLC mapping methodology and the consideration of enhancing that dataset with a people-centred, bottom-up approach for assessing the values of places in wider landscapes. What matters most to people in their living and working landscapes was captured as individually held values using social media data, online surveys, and semi-structured interviews. The dataset was subsequently categorised, following principles of *Grounded Theory* to allow themes to emerge from the stories, experiences, and connections of people to places. The categorisation compared the approach of a pre-defined and extended assessment system – Historic England's *Conservation Principles* and the narrative-based categorisation approach, where categories were developed on the emerging themes of the automated Topic Modelling process.

The classical Artificial Intelligence techniques adopted, adapted, and applied in this research – Natural Language Processing, Named Entity Recognition and Topic Modelling – provided the tools to approach qualitative data (unstructured text) through unsupervised learning. The advantage of the method is the ability to gain first insights into text-based data that is free of the researcher's assumptions and replicable, as this step in the analysis is based on statistical principles. This approach enabled the development of categories based on themes latent in or emerging from the empirical data.

GIS allowed querying the dataset based on specific questions and issues and provided maps for visual representation of social value patterns across landscapes that formed 'invisible' or 'hidden' value communities based on the same values (value-based) or on the same location (place-based).

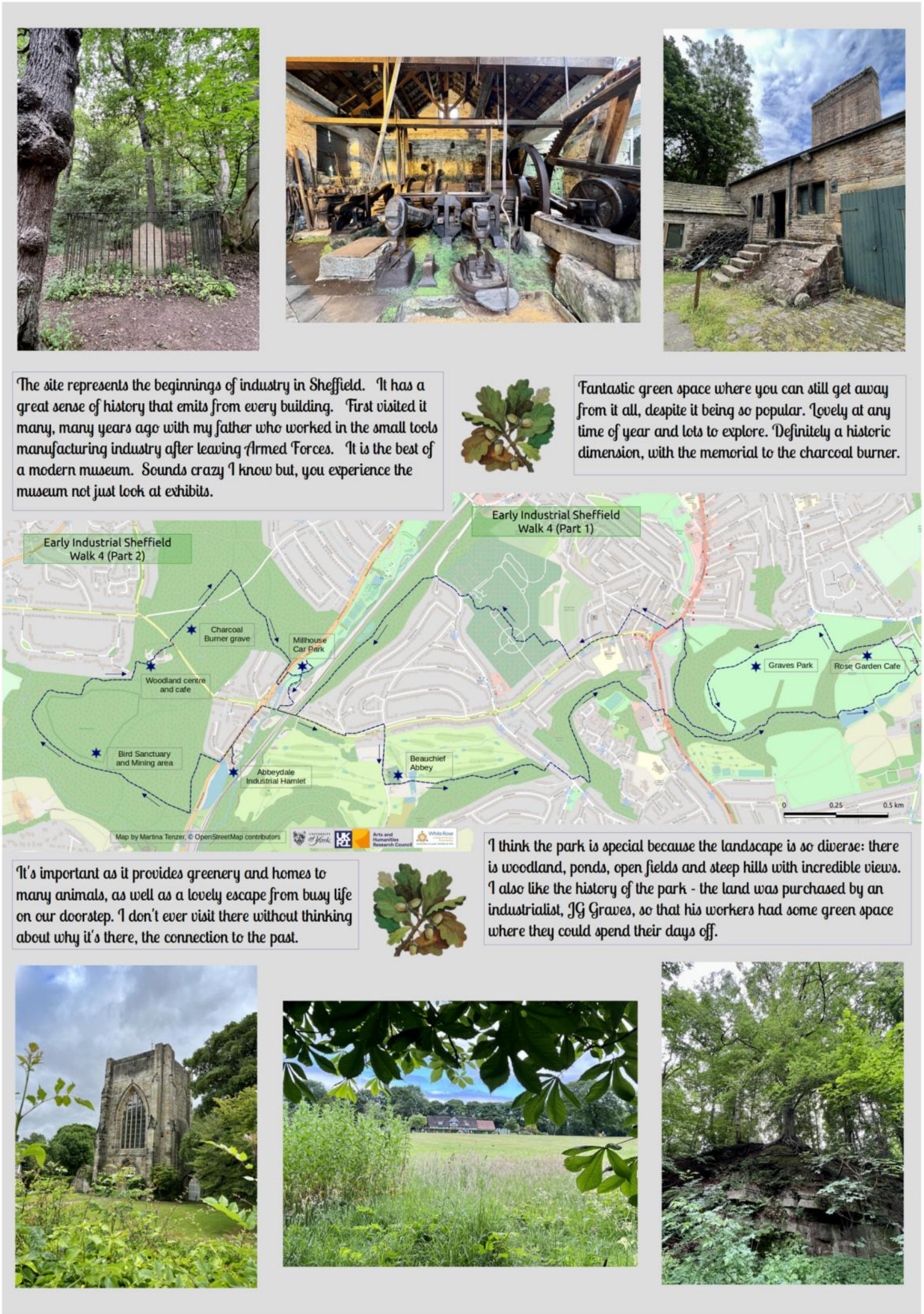

Figure 8-4: Leaflet for a self-guided walk presented at an outreach day of the PDNP, based on the stories and favourite locations of the survey participants.

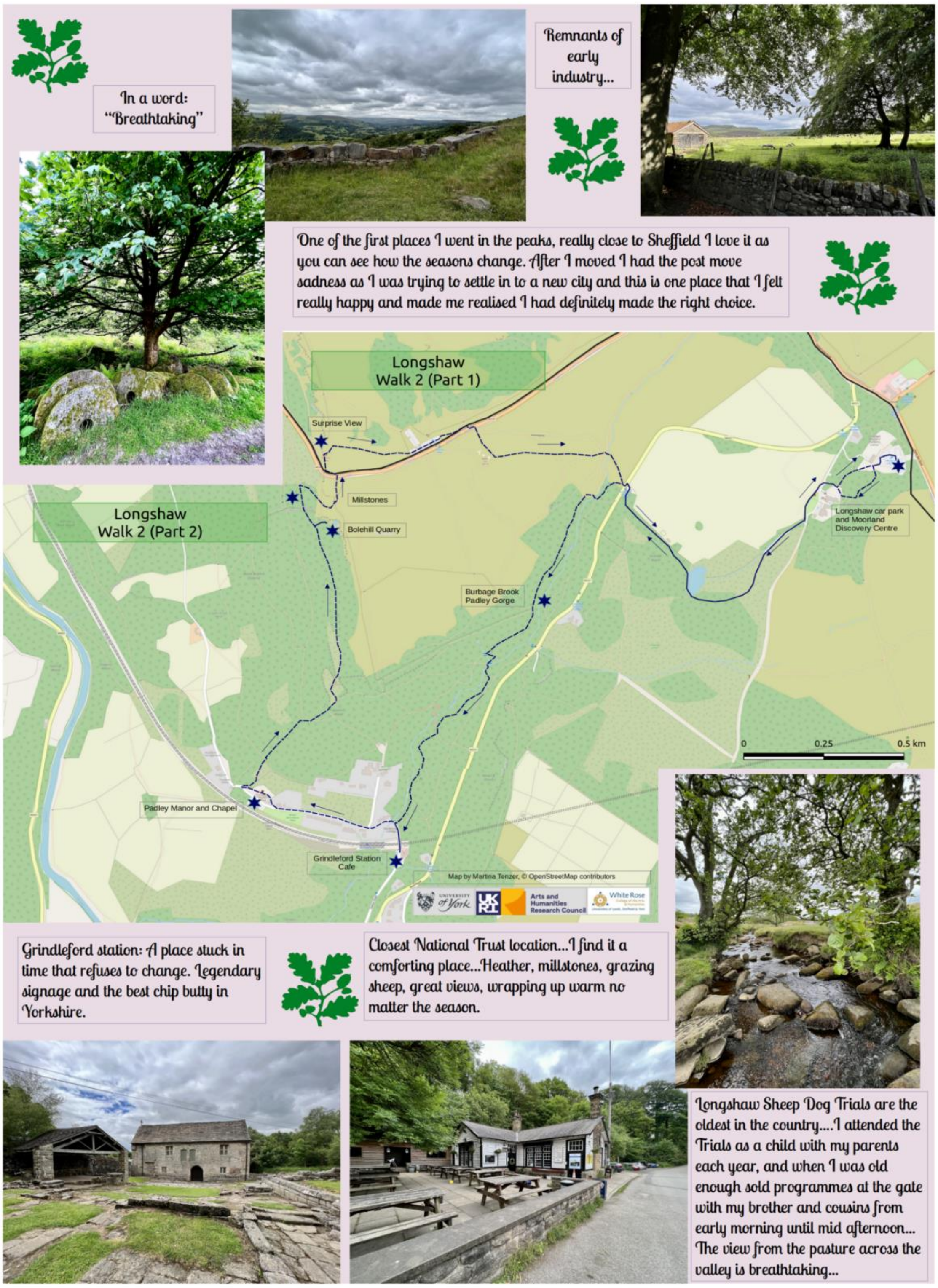

Figure 8-5: Leaflet of self-guided walk presented at an outreach day in the PDNP based on the stories of survey participants.

Existing tools and methods, as presented and discussed in **Chapter 1.2**. provided a basis for the identification of a gap in current approaches to social value assessment and opportunities to enhance and complement these. The overview showed that the method develop in this research can fill the gap to identify 'hidden' or 'invisible' communities based on the same social values and target these with existing tools for an in-depth analysis. Also, this method can provide a tool to identify hotspots of social value and target these for site-based analysis to help understand place attachment of communities. The dataset resulting from this research will enable a proactive engagement with sites and communities as opposed to the practice of reactive management to which current tools and methods provide solutions. One such existing tool for landscape assessment – HLC – was more closely assessed for compatibility. HLC was found to be a valuable tool for assessing historical landscapes but incapable of accommodating social values as an integral part of the dataset. However, *Social Landscape Characterisation*, conveying the needs, vision, experiences, perceptions, and aspirations of local people and local knowledge, provides a tool that is able to fulfil this function and integrate social value into the framework of official assessment strategies. Both datasets used in conjunction can enable informed and comprehensive planning and decision-making to facilitate inevitable change and development of landscapes. SLC offers a tool to enhance the understanding of place attachment, sense of place, identity and belonging. People-centred, place-based heritage and landscape management can increase the quality and distinctiveness of landscapes in an inclusive and transparent way and manage the historic environment socially sustainable for present and future generations.

8.7 Summary of ethical implications, bias, and limitations

The challenges for the methodology and techniques used in this research regarding ethical implications, bias and limitations were put into the wider context of AI applications in the discipline of archaeology and cultural heritage management in **Chapter 7**. Furthermore, **Chapters 3** to **5** detailed the specific implications and considerations for the specific approaches described for the applied techniques and created outputs of each data source. This section will summarise the implications and considerations to emphasise the responsible use of the methodology developed and applied in this thesis.

The ethical considerations should focus on the stakeholders regarding benefits and disadvantaging factors, such as exclusion, marginalisation, and bias towards specific groups in society. Potential stakeholders and applicants of the method include – but are not limited – to local authorities, local communities, and individuals with an interest in place attachment and sense of place (for example, for petitioning). Surveys and opinion polls are commonly used across local authorities to gauge sentiment and gather information about specific projects, for instance, for planning and development purposes. Introducing unconscious bias can negatively impact on specific groups within the community, limiting fairness and inclusion. Marginalised and vulnerable groups can be left out or disadvantaged in various stages of this method, for example, depending on the selection process of participants in the data collection and by automated processes, for example, by Artificial Intelligence training datasets (Casilli 2019; Raji et al. 2020). Decisions taken at every step of the process can have an impact on the output and, subsequently, on the people impacted by the resulting decision-making based on such outputs.

The methods and tools used in this research were subject to a rigorous internal ethical approval process at the University of York. However, the situation at the start of this thesis with the unfolding COVID-19 pandemic had an impact on and led to a redesign of the research methodology regarding the data collection and, subsequently, the data analysis (see **Chapter 1.5.2**). Ethical considerations, associated with the social media platform as a data source or the use of Topic Modelling were not part of the initial ethical considerations and approval process and required changes and amendments. To adhere to the restrictions around social contact during the COVID-19 pandemic, data collection had to be limited to remote and socially distant data collection methods. The commonly used data gathering methods, such as focus groups and in-person surveys and interviews were adapted to online surveys and interviews (Zoom, later in-person when restriction was eased). The use of social media was adopted later, and the ethical approval amended accordingly. The *Twitter Academic Developer Account*, as it was defined at the time of the research, allowed unlimited data access back to the beginnings of Twitter in 2006[25]. According to the terms and conditions of the platform and the ethical approval requirements of participants'

[25] The regulation of the platform now under the name of *X* has fundamentally changed since the time of writing the publication in 2022. The free access of the data is now not possible anymore and data access is limited.

anonymity, the data published and retained from this research includes only summaries, synthesised data, and statistics (**see Chapter 3**).

As mentioned in the introduction (**Chapter 1.5.4**), the dataset size changed because of the amended data collection methods. Instead of manual analysis or the use of NVivo, Artificial Intelligence techniques were adopted, adapted, and applied to further the understanding of AI in archaeology and heritage management. With over 2000 tweets and almost 500 stories of survey participants, the data sizes were smaller than commonly used data volumes in AI applications but too large to assess and evaluate manually in an efficient manner. The ethical approval was amended accordingly to address the implications for the privacy of participants and to minimise bias in the output of the analysis for management purposes (see **Appendices for Chapter 8/6**). The participants' consent was sought for data analysis but not for the use of the personal data as a basis for AI models, to be used for predictions and, subsequently, for shaping agendas or management decisions of local authorities (see **Appendices for Chapter 8/3** and **7**). Use of personal data for such purposes would have required an explicit and clear statement of this purpose in the consent form. This limited the scope of AI modelling work performed. The use of personal data for training models and the automatic categorisation of social values based on, for example, surveys would be an interesting field of investigation, particularly with the extension of the method in annual review cycles and with increasing dataset sizes. The use of methods such as Zero-shot or Few-shot learning[26], would have the potential for predictive social value categorisation. However, this approach was not covered by the participants' consent and, therefore, lay beyond the scope of this research. Nevertheless, following further research in a cross-disciplinary team these methods could provide a valuable tool for automated, time-efficient social value analysis.

[26] Zero- or Few-Shot learning are methods to use a machine learning model to predict new data based on either no labelled or only a small number of labelled data. Zero-shot learning makes predictions about new data based on relationships identified and characterised in data the model already knows. For Few-shot learning, the model predicts the category of the new data based on a few examples the model can learn from. Both methods are particulary useful for small datasets, as it was the case in this research.

When using AI techniques, the key principles underlying this research were: privacy, explainability and transparency[27]. Preserving privacy is an ongoing challenge for personal data in the public domain (Bender et al. 2021). Privacy in the data collection phase was guaranteed through the anonymisation of personal data and, where this was not possible, e.g., the interviews, the informed consent of participants included the waiving of anonymity (see **Appendix for Chapter 8/5**). As mentioned, a later use of the data for modelling purposes was not possible, as this was not set out in the consent signed by the participants. Explainability during the analysis phase meant that the 'black box' effect had to be minimised as much as possible. Deterministic, statistical AI techniques were used and executed on an automated basis, while most steps in the workflow included a human element of assessment and parameter setting (Chang et al. 2009; Lau, Newman and Baldwin 2014). Therefore, the processing of data remained comprehensive and replicable[28]. Transparency, as part of the output process, guaranteed that methods and tools, as well as outputs, were understood against the background of introduced bias and limitations of the method. Limitations and bias of the specific data sources and methods used for this analysis were highlighted in **Chapters 3 to 5**. They included self-selected participants through means of survey (online), members of the mailing list of Sheffield City Council, users of social media for the publication of the PDNP survey and social media analysis, convenience sample method and typical case sample method for interviews (see **Chapter 1.5.5**). These factors, particularly during the data collection process, excluded potential participants with no access to social media or the internet. Therefore, it was essential to raise awareness for the people left out or unfairly treated or assessed during any analysis stage[29]. Further work could potentially include specifically marginalised or vulnerable people in the analysis.

---

[27] For more on ethical implication, bias and limitations of AI technology see **Chapter 7**.

[28] The modelling phase in this research was based exclusively on statistical methods (Latent Dirichlet Allocation, Gibbs Sampling). Bias can be introduced where training data is used on datasets with varying quality and content.

[29] See **Chapter 3**, Footnote 4 for the use of additional general statistics inferred from the survey responses. The significant number of 91.6 per cent identified themselves as White British in the

Furthermore, while the method based on *Grounded Theory* (Charmaz 2006; Odacioglu and Zhang 2022) and statistical analysis methods reduced the researcher bias to a certain degree, subjectivity was introduced during the manual assessment or direct observation process. Nevertheless, using AI techniques, particularly TM, to explore latent themes within empirical data allowed approaching stories of participants without predefined categories, codes, or assumptions and preconceptions. Qualitative data was approached from a new angle to reveal latent themes emerging from empirical data. Furthermore, the automated analysis ensures replicability and reproducibility of the process, which is commonly inconsistent and subjective if carried out by human assessors, even if codes and approaches to data are agreed upon.

As the main output of this thesis is presented in a map format, it is important to note that maps have the potential to be misused or misinterpreted (Monmonier 1996). Maps are specifically well suited to convey complex or landscape-wide information. However, they have limits, for instance, what questions can be answered with a single output. Multi-layered GIS maps have overcome this problem by allowing filtering to specific datasets or layers. However, as with HLC maps, the content and amount of information can only answer clearly stated questions. Therefore, maps need a specifically defined purpose and should only be used and interpreted based on this premise. If used according to the intended purpose, maps are ideal for visualising and presenting data in an accessible and scalable format.

Maps created in this research varied from highly complex maps (e.g., HLC maps) to engaging maps for public engagement (emoji maps). For example, the emoji map (see **Chapter 3**) was intended to show the frequency of emojis used in tweets about the PDNP. To visualise this emoji cloud, the outline of the national park was used as a basic shape and populated with emojis of varying sizes according to the frequency of use, similar to word clouds. However, this design has the potential to give the false impression that the emojis are exactly spatially located on the map of the PDNP. This impression of geographic location is supported by using a map outline as a shape (instead of a circle or square, which would

survey, revealing shortcomings of the distribution channels and technology used (e.g., social media, computer, and internet access).

fulfil the same purpose). However, this is not the case, as the frequency cannot simultaneously show the position of the emojis used. This would only be possible if, for example, the focus would have been on one type of emoji, showing its distribution across the national park. Frequency and distribution cannot be visualised at the same time for different types of data. Such pitfalls and shortcomings tend to be misunderstood or misinterpreted and should be preempted and tested at the outset of a project.

In summary, digital and AI techniques enable researchers to analyse larger datasets more time-efficiently and approach, link and understand data from new angles. However, without a 'human in the loop' there is an inherent risk of enforcing unrecognised biases in algorithms or training data. The techniques used in this research also brought new challenges and implications to ensure privacy during the data collection process, explainability in the analysis phase and transparency of the final output. Technologies are fast evolving, and research can not only benefit from the opportunities but must also ensure safe, responsible, just, and inclusive handling of data and use of outputs. Nevertheless, the method presented in this research has proven effective, efficient, and a contribution to the evolving field of AI applications in qualitative research of archaeology and CHM.

8.8 Personal reflections on challenges and successes

The research for this thesis depended on various factors, some of which were anticipated in advance and others that evolved over the course of the project. Looking back on the challenges and success of this project, I want to detail the most important factors that shaped and determined this research.

Firstly, the partnership with local authorities and dependence on public participation, as well as the handling of personal data posed challenges and advantages as an essential part of this research. Data collection, for instance, depended on disseminating the online questionnaire through the media channels of my partner organisations. Not having full control over this part of the project necessitated more flexibility in the planning and accepting delays of survey responses for further analysis. On the other hand, the partnership with local authorities also opened new opportunities for disseminating my online surveys. It introduced me to interview partners, who would otherwise have been more difficult to reach. The significant number of participants in my survey was undoubtedly a successful outcome of this partnership. With almost 500 respondents to my online questionnaire, the project established a database that allowed sufficient coverage of the study areas.

Secondly, the use of social media as a data source for the analysis was an element of my research, which developed later in the project to address the restrictions of the COVID-19 pandemic on personal contact with participants for data collection. The analysis of social media data provided a treasure trove of information on the PDNP and sentiment towards the landscape at different points in time. At the time of writing the publication for this research, access to the data was free and subject to a *Twitter Academic Developer Account*, and I considered Twitter to be the most researcher-friendly social media platform. This approach gave me a broad overview of trending topics (hashtags) and emotions associated with my study area across different years. I saw this as a successful part of my research. However, the fast-evolving events in the social media landscape, and, particularly with the changed ownership of Twitter and, subsequently, the changed terms and conditions of use, made my research approach unfeasible for future research. Access to the data has been restricted and is now highly charged, on the one hand with high fees and, on the other hand, with claims of increased misinformation disseminated through X, as the platform is now known. The data for research is also restricted and does not allow free access to all tweets. Therefore, at the time of writing, X is no longer suitable for a research approach as conducted as part of this project.

Rapid change and development were also challenges for the use of Artificial Intelligence as a technique to analyse the data. At the outset of the project, I was unfamiliar with AI techniques, such as Natural Language Processing (NLP) or Topic Modelling (TM). Developing my coding and AI skills was a steep, but enriching learning curve. However, similar to social media, the environment of digital technologies and AI is fast evolving. The developments during the years of this research posed opportunities but also challenges. I only scratched the surface of the possibilities this technology has to offer. At times, I applied a technique with implications that only revealed themselves afterwards. This meant carefully considering the ethical implications, limitations and introduced bias to reflect on the outcome of my analyses before I was able to publish my research. The outcome of this deep immersion into the topic of AI ethics resulted in me organising a conference on AI ethics in academic research and a publication on AI applications and implications in the discipline (see **Chapter 7**). Nevertheless, the outcome of the AI applications is a success of this thesis by treading on fairly new and unfamiliar ground for the analysis of text-based data in archaeology and Cultural Heritage Management. In the meantime, new AI techniques, like Large Language Models (LLM), in particular Generative Pre-trained Transformers (GPT) have overshadowed classical NLP techniques and TM as used in this thesis. However, while these new techniques show promise in the application discussed in this work, their black box character, and thus their non-explainability, poses risks that have been discussed in detail in **Chapter 7**. The classical NLP and TM techniques used in this project are, therefore, still a valid and safe approach, while keeping the option for a drop-in replacement with GPT based on Zero- or Few-Shot methods, once the risks and implications of the GPT method, which are only just coming to light, are better understood.

Finally, opting for a thesis by publication and focusing on the publication of articles throughout this research was a challenge and benefit in several ways. Choosing this thesis format allowed me to publish my results in a timely manner instead of waiting for the completion of the PhD. Topics such as Twitter analysis or AI ethics discussion, are important elements of this research. However, as mentioned before, immediate publication of my results contributed to the current research environment and enhanced my research through feedback from reviews while, at the same time, increasing my resilience, appreciation of and constructive approach to critique. Furthermore, the experience of publishing a series of articles contributed to my personal and professional development as an independent researcher. The publication process was, again, outside of my control, with sometimes long processing time, and required additional flexibility in my planning. Nevertheless, the experience of the publication process was certainly a benefit.

This project was based on exploring new technologies and experiences that offered opportunities and introduced challenges. The successful conclusion of this project offers an innovative approach and raises awareness for the application of AI technology. At the same time, this research opened new opportunities for further research and development.

8.9 Conclusion and opportunities for further work

8.9.1 Conclusion

The approach taken in this research addresses the challenge to collect, analyse and meaningfully integrate individually held values into the assessment framework of heritage and landscape management. Including individually held values into systematic assessment and rigid categorisation has previously been seen as too challenging due to the dynamic and fluid nature of values and the non-conformity of individual opinions.

This thesis provides an innovative approach to qualitative data analysis in heritage and landscape studies to assess and categorise individually held values across wider natural and cultural landscapes with the potential for continuous assessment. This approach offers opportunities for engagement and participation of the public and provides an essential resource as background for planning decisions within local authorities to facilitate socially and environmentally sustainable change and development. Key findings of this thesis were the meaningful analysis of individually held values in the historic environment, a methodology for enabling the dynamic and fluid categorisation of social values, and the visual representation of social values, while addressing their dynamic nature of social values.

Firstly, the methodology, based on Artificial Intelligence techniques allowed a time- and resource-efficient method to collect and analyse individually held values across landscapes, based on the personal stories of people. These stories encompassed personal family history, a connection to the local community, traditions, events, local history, as well as the grand and designated national important heritage assets. The tangible and intangible elements of the living and working landscapes can form an essential background of insider knowledge or local expertise to aid informed decision-making for socially sustainable change and development of heritage and cultural landscapes.

Secondly, exploring the merits of categorisation frameworks revealed that, on the one hand, an existing system like the *Conservation Principles* can be adapted to accommodate the needs, visions and aspirations of people expressed in their personal stories of connection to places. On the other hand, free, narrative-based categorisation showed more variety and flexibility of the categories which emerged from the stories. However, both these categorisation systems showed high levels of congruence. In both cases, Topic Modelling, as a method to capture the essence of people's perceptions, experiences, and connections, provided an innovative approach to narrative-based categorisation. A value-based landscape assessment, derived from categorisation, was necessary for visualising individually held values and the meaningful integration in GIS for use in official assessment frameworks.

Thirdly, this thesis offers examples of the visual representation of social values, which form patterns of previously unknown, 'hidden' or 'invisible' values held by people living and working in these landscapes. Patterns emerged based on the same values across areas (value-based), derived from the categorisation, or valuing the same place for different reasons (place-based).

Furthermore, the study provides a methodology for the analysis of data in the form of a lab book and a detailed methodology, including the code for analysis, to repeat and recreate the findings in varying scales, areas and focus groups (for the detailed methodology and lab book see **Appendices for Chapter 8/1a and 1b**). Based on freely available open-source software and code, the method can be applied by local authorities, community and interest groups or individual people interested in exploring individually held and social values embedded in their neighbourhoods, areas, or wider landscapes[30].

Social Landscape Characterisation (SLC) provides an essential and replicable framework to incorporate locally and individually held social values into the assessment practice of local authorities, enabling proactive and collaborative work with people living and working in these landscapes that are dynamic and ever-changing. Landscapes consist of designated, locally listed, or everyday heritage and natural or semi-natural features, which are perceived and valued by people. Integrating people's perceptions, experiences, needs, visions, and aspiration regarding place into official assessment frameworks and using the dataset in conjunction with HLC, SLC can offer a tool to enhance the understanding of place attachment, sense of place, identity and belonging. People-centred, place-based heritage and landscape management can increase the quality and distinctiveness of landscapes in an

[30] As the code is provided in the lab book format, a basic knowledge of coding is necessary to apply the code. See also **Chapter 8.9.2**.

inclusive and transparent way and manage the historic environment in a socially sustainable way for present and future generations. As discussed in **Chapter 1.2**. current approaches focus on specific aspects of place attachment, using quantitative (Likert Scales) or structured interviews and surveys and, thereby, do not offer the open and narrative approach as taken in this research (see **Chapter 1.2.4**). Also, social value and community toolkits and methods are particularly focussed on pre-defined communities or local groups and cannot provide an individually held knowledge base to reveal 'hidden' or 'invisible' value communities (see **Chapter 1.2.7**). Finally, current approaches of social values assessment focus on projects or sites in a reactive approach and cannot provide information of independently held values across wider landscapes to form a better understanding of social values for the purpose of proactive planning and decision-making.

This thesis provides a methodology for the assessment and integration of individually held values which form patterns of social values across wider landscapes. At the same time, this research also opens new opportunities for further research. Opportunities for inter- and transdisciplinary research will be detailed in the final section of this thesis.

8.9.2 Further work

This research has included elements of different disciplines, including Cultural Heritage Management (CHM) and Archaeology, Urban Planning, Human Geography, Anthropology, Social Science and Data Science. The exploration and integration of tools and techniques from these disciplines opened potentially new opportunities for further inter-, cross- and trans-disciplinary research in the field of natural and cultural landscapes and heritage management.

The code provided in the lab book accompanying this thesis remains in the raw R code version, as the development of an end-user application lies beyond the scope of this thesis due to time constraints. However, developing a user-friendly GUI for an application of the method could be part of a master's dissertation in Data or Computer Science. This would enable people with no or little knowledge in coding to use the analysis method on their collected data or stories.

During this research, the application of AI techniques was adopted and adapted to model topics in qualitative data. The next step for the automated categorisation would be to train a model on assessed and categorised data and apply Zero-shot and Few-shot learning for predictive analysis and categorisation of qualitative data. This would speed up the categorisation of new data which was anticipated to be collected on an annual or continuous basis. This would allow to address the fluid and dynamic nature of social values and review theses on an annual basis. This method would provide local authorities with an up-to-date database of social values for various purposes, for example, planning and decision making, and outreach and engagement. This research would benefit from exploring the method, in general, with larger datasets which can be collected on a rolling basis. Larger datasets would benefit the extension of the database across the landscape and minimise the blank areas with missing data for social values and would allow a more effective training dataset for the predictive categorisation of new data.

Anthropology and Social Sciences may use qualitative data for further place attachment research. Research that will become increasingly important in the field of urban planning for place-making and public benefit. Social sciences could explore the potential of this method in form of applications for public engagement and inclusion in various sectors expanding the application to health and wellbeing, inclusion of marginalised and vulnerable groups and effects of social injustice.

Archaeology and CHM are only slowly adopting and adapting AI techniques and tools for the analysis of data. While the application of AI in the field of image analysis is more advanced, the text-based analysis of archaeological or heritage data is slowly adopting AI. For further analysis and correlation of heritage datasets, for example, HLC and LCA datasets an extension of analysis, practical application and tests in real-world scenarios would benefit the development of the method developed in this research.

For the limitations of this research, it would also be a field of further potential work to include marginalised or vulnerable groups in the methodology and addressing social values against the background of special needs and particular situations of parts of society.

8.10 Future application of the methodology

The methodology and code were presented to the partners of this project – the Peak District National Park and the city of Sheffield – for testing and application to real-world scenarios. An annual review and extending dataset would have the potential to provide a database of social values that can help the PDNPA and Sheffield City Council gauge people's emotional attachment to place and manage landscapes accordingly.

Beyond the study areas, the importance of social values and people-centred, place-based approach to assess places and landscapes in the planning stage is evidenced by the national and international interest in SLC as a reaction to the publications. The method was presented to Historic England, with a potential to contribute to the corporate plan to foster people's 'pride in their local place' (Priority 1, *Levelling Up*) and to increase inclusion (Priority 2: *Inclusion, Diversity and Equality*)(Historic England 2023).

Also, an invitation to present SLC to Australian heritage consultancies (e.g., EMM Consulting Pty Ltd., Sydney) shows that the global phenomenon of public engagement and participation has gained a vital role in place-making and public benefits in heritage and landscape management.

The genuinely democratic approach presented by SLC will enable communication, cooperation and collaboration between professionals and laypeople in a bottom-up approach that has been seen as too challenging in the past. *Social Landscape Characterisation*, as a method based on the needs, visions, and aspirations of the residents for their places, can help heritage managers understand how to facilitate change and development in a socially sustainable way and, at the same time, foster the appreciation for heritage and reinforce a sense of place, belonging and identity for every individual person.

References

Abram, M. D., Mancini, K. T. and Parker, R. D. (2020). Methods to integrate natural language processing into qualitative research. *International Journal of Qualitative Methods*, 19, p.1609406920984608. [Online]. Available at: doi:10.1177/1609406920984608.

Ahmer, C. (2020). Riegl's 'Modern cult of monuments' as a theory underpinning practical conservation and restoration work. *Journal of Architectural Conservation*, 26 (2), pp.150–165. [Online]. Available at: doi:10.1080/13556207.2020.1738727.

Aldred, O. and Fairclough, G. (2003). *Historic Landscape Characterisation - taking stock of the method | Historic England*. [Online]. Available at: http://historicengland.org.uk/images-books/publications/hlc-taking-stock-of-the-method/ [Accessed 29 September 2020].

Altman, I. and Low, S. M. (Eds). (1992). *Place attachment*. New York: Plenum Press.

Anichini, F. et al. (2021). The automatic recognition of ceramics from only one photo: The ArchAIDE app. *Journal of Archaeological Science: Reports*, 36, p.102788. [Online]. Available at: doi:10.1016/j.jasrep.2020.102788.

Argyrou, A. and Agapiou, A. (2022). A review of artificial intelligence and remote sensing for archaeological research. *Remote Sensing*, 14 (23), p.6000. [Online]. Available at: doi:10.3390/rs14236000.

Arnstein, S. R. (2019). A ladder of citizen participation. *Journal of the American Planning Association*, 85 (1), pp.24–34. [Online]. Available at: doi:10.1080/01944363.2018.1559388.

Australian Heritage Council. (2009). *Guidelines for the assessment of places for the National Heritage List*. [Online]. Available at: https://www.dcceew.gov.au/parks-heritage/heritage/ahc/publications/nhl-guidelines [Accessed 30 September 2023].

Avrami, E. (2009). Heritage, values and sustainablitiy. In: Richmond, A., Bracker, A. L. and Victoria and Albert Museum (Eds). *Conservation: principles, dilemmas and uncomfortable truths*. 1st ed. Oxford, Amsterdam : Butterworth-Heinemann, Elsevier/Butterworth-Heinemann in association with the Victoria and Albert Museum London. pp.177–183. [Online]. Available at: https://ebookcentral.proquest.com/lib/york-ebooks/detail.action?docID=535314 [Accessed 4 February 2021].

Avrami, E. and Mason, R. (2019). Mapping the issue of values. In: Macdonald, S. (Ed). *Values in heritage management: emerging approaches and research directions*. Los Angeles: Getty Publications. pp.9–33.

Avrami, E., Mason, R. and de la Torre, M. (2000). *Values and heritage conservation*. Los Angeles : Getty Conservation Institute. [Online]. Available at: http://www.tandfonline.com/doi/full/10.1179/2159032X13Z.00000000011 [Accessed 16 November 2020].

Bailey, J. and Biggs, I. (2012). “Either side of delphy bridge”: a deep mapping project evoking and engaging the lives of older adults in rural North Cornwall. *Journal of Rural Studies*, 28 (4), pp.318–328. [Online]. Available at: doi:10.1016/j.jrurstud.2012.01.001.

Baker, C. et al. (2021). *Coronavirus: a history of English lockdown laws*. [Online]. Available at: https://commonslibrary.parliament.uk/research-briefings/cbp-9068/ [Accessed 26 January 2022].

Banks, G. C. et al. (2018). A review of best practice recommendations for text analysis in R (and a User-Friendly App). *Journal of Business and Psychology*, 33 (4), pp.445–459. [Online]. Available at: doi:10.1007/s10869-017-9528-3.

Barai, M. K. (2021). Sentiment analysis with textblob and vader in python. *Analytics Vidhya*. [Online]. Available at: https://www.analyticsvidhya.com/blog/2021/10/sentiment-analysis-with-textblob-and-vader/ [Accessed 11 February 2022].

Barnatt, J. (2003). *A landscape through time. The historic character of the Peak District National Park landscape. Aims, methods and user manual.* Bakewell: Peak District National Park Authority. [Accessed 14 December 2017].

Barnatt, J. and Penny, R. (2004). *The distribution of lead mining surface remains in the Peak District*. Bakewell : Peak District National Park Authority. [Accessed 10 February 2019].

Barrie, C. (2022). *academictwitteR*. R. [Online]. Available at: https://github.com/cjbarrie/academictwitteR [Accessed 12 January 2022].

Barrie, C. and Ho, J. C. (2021). academictwitteR: an R package to access the Twitter Academic Research Product Track v2 API endpoint. *Journal of Open Source Software*, 6 (62), p.3272. [Online]. Available at: doi:10.21105/joss.03272.

Bell, S. (2005). In: Kowarik, I. and Körner, S. (Eds). *Nature for people: The importance of green spaces to communities in the East Midlands of England*. Heidelberg: Springer. pp.81–94. [Online]. Available at: https://www.academia.edu/69012686/Nature_for_People_The_Importance_of_Green_Spaces_to_Communities_in_the_East_Midlands_of_England [Accessed 16 January 2023].

Bender, E. M. et al. (2021a). On the dangers of stochastic parrots: can language models be too big? 🦜. In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. FAccT ’21. 1 March 2021. New York, NY, USA : Association for Computing Machinery. pp.610–623. [Online]. Available at: doi:10.1145/3442188.3445922 [Accessed 16 May 2023].

Benesch, S. (2021). *Nobody can see into Facebook*. [Online]. Available at: https://www.theatlantic.com/ideas/archive/2021/10/facebook-oversight-data-independent-research/620557/ [Accessed 11 February 2022].

Bertrand, K. Z. et al. (2013). Sentiment in New York City: a high resolution spatial and temporal view. *arXiv:1308.5010 [physics]*. [Online]. Available at: http://arxiv.org/abs/1308.5010 [Accessed 28 August 2021].

Betjeman, J. author. (1947). *Slick but not streamlined: poems & short pieces / by John Betjeman ; selected, & with an introduction by W.H. Auden.* Garden City, N.Y.: Doubleday & Co.

Bevan, A. (2015). The data deluge. *Antiquity*, 89 (348), pp.1473–1484. [Online]. Available at: doi:10.15184/aqy.2015.102.

Bianchini, R. (2021). *Museums worldwide react to COVID lockdown by offering virtual tours | Inexhibit*. [Online]. Available at: https://www.inexhibit.com/marker/museums-worldwide-react-to-covid-lockdown-by-offering-virtual-visits/ [Accessed 11 February 2022].

Bickler, S. H. (2021). Machine learning arrives in archaeology. *Advances in Archaeological Practice*, 9 (2), pp.186–191. [Online]. Available at: doi:10.1017/aap.2021.6.

Blei, D. M., Ng, A. Y. and Jordan, M. I. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research*, 3 (0), pp.993–1022.

Bodenhamer, D. J. et al. (2015). *Deep maps andsSpatial narratives*. Bloomington: Indiana University Press. [Online]. [Accessed 8 January 2021].

Boley, B. B. et al. (2021). Measuring place attachment with the Abbreviated Place Attachment Scale (APAS). *Journal of Environmental Psychology*, 74, p.101577. [Online]. Available at: doi:10.1016/j.jenvp.2021.101577.

Bonnell, V. E. and Hunt, L. A. (1999). Introduction. In: *Beyond the cultural turn: new directions in the study of society and culture*. ACLS Humanities E-Book. Berkeley, Calif., Berkeley, Calif.; London: University of California Press. pp.1–32. [Online]. Available at: http://hdl.handle.net/2027/heb.04780 [Accessed 19 January 2021].

Bordoni, L., Mele, F. and Sorgente, A. (Eds). (2016). *Artificial intelligence for cultural heritage - Cambridge Scholars Publishing*. Cambridge: Cambridge Scholars Publishing. [Online]. Available at: https://www.cambridgescholars.com/product/978-1-4438-9085-4.

Bowden, M. (2001). Mapping the past: O. G. S. Crawford and the development of landscape studies. *Landscapes*, 2 (2), pp.29–45. [Online]. Available at: doi:10.1179/lan.2001.2.2.29.

Bradley, A. et al. (2004). *Change and creation: historic landscape character 1950-2000 | Historic England*. London: English Heritage. [Online]. Available at: http://historicengland.org.uk/images-books/publications/change-and-creation-historic-landscape-character/ [Accessed 9 December 2022].

Brandsen, A. and Lippok, F. (2021). A burning question – using an intelligent grey literature search engine to change our views on early medieval burial practices in the Netherlands. *Journal of Archaeological Science*, 133, p.105456. [Online]. Available at: doi:10.1016/j.jas.2021.105456.

Bristol City Council. (2021). *Know your place - bristol.gov.uk*. [Online]. Available at: https://www.bristol.gov.uk/planning-and-building-regulations/know-your-place [Accessed 14 February 2021].

Brotton, J. (2015). *Great maps: the world's masterpieces explored and explained*. London: Penguin Random House.

Brown, G. (2004). Mapping spatial attributes in survey research for natural resource management: methods and applications. *Society & Natural Resources*, 18 (1), pp.17–39. [Online]. Available at: doi:10.1080/08941920590881853.

Brown, G. and Raymond, C. (2007). The relationship between place attachment and landscape values: toward mapping placeaAttachment. *Applied Geography*, 27, pp.89–111. [Online]. Available at: doi:10.1016/j.apgeog.2006.11.002.

Brown, G., Raymond, C. M. and Corcoran, J. (2015). Mapping and measuring place attachment. *Applied Geography*, 57, pp.42–53. [Online]. Available at: doi:10.1016/j.apgeog.2014.12.011.

Brown, G. and Weber, D. (2011). Public Participation GIS: A new method for national park planning. *Landscape and Urban Planning*, 102 (1), pp.1–15. [Online]. Available at: doi:10.1016/j.landurbplan.2011.03.003.

Brown, J., Mitchell, N. and Beresford, M. (Eds). (2005). *The protected landscape approach: linking nature, culture and community*. Gland, Switzerland : IUCUN. [Online]. Available at: https://portals.iucn.org/library/sites/library/files/documents/2005-006.pdf [Accessed 26 December 2020].

Bryant, J. C. and Farrell, J. (2023). The influence of the nature-culture dualism on morality. In: Hitlin, S., Dromi, S. M. and Luft, A. (Eds). *Handbook of the sociology of morality, Volume 2*. Handbooks of Sociology and Social Research. Cham: Springer International Publishing. pp.261–275. [Online]. Available at: doi:10.1007/978-3-031-32022-4_18 [Accessed 26 December 2023].

Butler, A. (2016). Dynamics of integrating landscape values in landscape character assessment: the hidden dominance of the objective outsider. *Landscape Research*, 41 (2), pp.239–252. [Online]. Available at: doi:10.1080/01426397.2015.1135315.

Byrne, D. (2008a). Countermapping: New South Wales and Southeast Asia. *Transforming Cultures eJournal*, 3 (1). [Online]. Available at: doi:10.5130/tfc.v3i1.687 [Accessed 23 December 2020].

Byrne, D. (2008b). Heritage as social action. In: Fairclouch, G. et al. (Eds). *The Heritage Reader*. Abingdon : Routledge. pp.149–173.

Byrne, D. (2014). Counter-mapping and migrancy on the Georges River. In: Schofield, J. (Ed). *Who needs Experts? Counter-mapping cultural heritage*. Abingdon: Routledge. pp.77–92.

Byrne, D., Brayshaw, H. and Ireland, T. (2003). *Social significance: a discussion paper*. 2nd ed. Hurstville, N.S.W. : NSW National Parks and Wildlife Service. [Online]. Available at: https://www.environment.nsw.gov.au/-/media/OEH/Corporate-Site/Documents/Aboriginal-cultural-heritage/social-significance-a-discussion-paper-010001.pdf.

Byrne, D. and Nugent, M. (2004). *Mapping attachment: A spatial approach to aboriginal post-contact heritage*. Hurstville: Department of Environment and Conservation.

Byrne, D. and Ween, G. B. (2015). Bridging cultural and natural heritage. In: Meskell, L. (Ed). *Global heritage: a reader*. Blackwell readers in anthropology. 1st ed. Somerset: John Wiley & Sons, Inc, John Wiley & Sons, Incorporated, Wiley-Blackwell. pp.94–111.

Cai, Z. et al. (2021). Using topic modelling for code discovery in large scale text data. In: Ruis, A. R. and Lee, S. B. (Eds). *Advances in quantitative ethnography*. Communications in Computer and Information Science. 2021. Cham: Springer International Publishing. pp.18–31. [Online]. Available at: doi:10.1007/978-3-030-67788-6_2.

Campaign to Protect Rural England. (2018a). *Unlocking the landscape*. [Online]. Available at: https://www.cpre.org.uk/wp-content/uploads/2019/11/unlocking_the_landscape1_1.pdf [Accessed 8 February 2021].

Campaign to Protect Rural England. (2018b). *What's special to you: landscape issues in your neighbourhood plan*. [Online]. Available at: https://www.cpre.org.uk/wp-content/uploads/2019/11/CPREZLandscapeZZNeighbourhoodZPlanningZtextZonly.pdf [Accessed 8 February 2021].

Cao, N. and Cui, W. (2016). *Introduction to text visualization*. Paris : Atlantis Press. [Online]. Available at: doi:10.2991/978-94-6239-186-4 [Accessed 13 February 2022].

Casilli, A. (2019). *En attendant les robots: enquête sur le travail du clic*. Paris : Seuil.

Casini, L. et al. (2023). A human–AI collaboration workflow for archaeological sites detection. *Scientific Reports*, 13 (1), Nature Publishing Group., p.8699. [Online]. Available at: doi:10.1038/s41598-023-36015-5.

Chang, J. et al. (2009). Reading tea leaves: how humans interpret topic models. In: Bengio, Y. et al. (Eds). *Advances in neural information processing systems*. 22. 2009. Curran Associates, Inc. [Online]. Available at: https://proceedings.neurips.cc/paper/2009/file/f92586a25bb3145facd64ab20fd554ff-Paper.pdf.

Charmaz, K. (2006). *Constructing grounded theory: a practical guide through qualitative analysis / Kathy Charmaz.* Los Angeles; London: Sage Publications.

Chitty, G. (2016). Introduction. In: *Heritage, conservation and communities: engagement, participation and capacity building*. Heritage, culture, and identity. London: Routledge. pp.1–14.

Chun Tie, Y., Birks, M. and Francis, K. (2019). Grounded theory research: a design framework for novice researchers. *SAGE Open Medicine*, 7, p.2050312118822927. [Online]. Available at: doi:10.1177/2050312118822927.

Cinderby, S. et al. (2012). Analyzing perceptions of inequalities in rural areas of England using a mixed-methods approach. *URISA, Journal of the Urban and Regional Information Systems Association*, 24 (2), pp.33–42.

Cinderby, S. and Forrester, J. (2005). Facilitating the local governance of air pollution using GIS for participation. *Applied Geography*, 25 (2), pp.143–158. [Online]. Available at: doi:10.1016/j.apgeog.2005.03.003.

Cinderby, S., Snell, C. and Forrester, J. (2008). Participatory GIS and its application in governance: the example of air quality and the implications for noise pollution. *Local Environment*, 13 (4), pp.309–320. [Online]. Available at: doi:10.1080/13549830701803265.

City of Lincoln Council. (2011). *Heritage connect Lincoln*. [Online]. Available at: http://www.heritageconnectlincoln.com/article/about [Accessed 14 February 2021].

Clark, J. (2003). *HLC applications review. Technical Report*. Preston: Lancashire County Council.

Clark, J., Darlington, J. and Fairclough, G. (2004). *Using historic landscape characterisation. English Heritage's review of HLC Applications 2002 - 03*. London and Preston: English Heritage & Lancashire County Council.

Clavert, F. and Gensburger, S. (2023). *Is artificial intelligence the future of collective memory? Bridging AI scholarship and memory studies*. [Online]. Available at: https://www.c2dh.uni.lu/sites/default/files/cfp_is_artificial_intelligence_the_future_of_collective_memory_-_2023.pdf.

Clifford, S. and King, A. (Eds). (1996). *from place to PLACE - maps and parish maps*. London : Common Ground. [Online]. Available at: https://www.commonground.org.uk/shop/from-place-to-place/ [Accessed 6 December 2020].

Cobb, P. (2023). Large language models and Generative AI, Oh My!: Archaeology in the time of ChatGPT, Midjourney, and Beyond. *Advances in Archaeological Practice*, 11, pp.363–369. [Online]. Available at: doi:10.1017/aap.2023.20.

Cohen, A., Klassen, S. and Evans, D. (2020). *Ethics in archaeological lidar*. 3 (1), Ubiquity Press., pp.76–91. [Online]. Available at: doi:10.5334/jcaa.48.

Common Ground. (1996). *Parish maps*. [Online]. Available at: https://www.commonground.org.uk/parish-maps/ [Accessed 25 September 2020].

Common Ground. (2006). Local distinctiveness. *Common Ground*. [Online]. Available at: https://www.commonground.org.uk/local-distinctiveness/ [Accessed 1 February 2021].

Common Ground. (2019). History. *Common Ground*. [Online]. Available at: https://www.commonground.org.uk/history/ [Accessed 14 February 2020].

Condorelli, F. et al. (2020). A neural networks approach to detecting lost heritage in historical video. *ISPRS International Journal of Geo-Information*, 9 (5), p.297. [Online]. Available at: doi:10.3390/ijgi9050297.

Cook, I. and Taylor, K. (2013). *A contemporary guide to cultural mapping. An ASEAN-Australia perspective*. Jakarta: ASEAN Secretariat.

Cosgrove, D. (2004). Lanscape and Landschaft. *Geman Historical Institute*, 35 (Fall), pp.57–71.

Council of Europe. (1975). *European Charter of the architectural heritage*. [Online]. Available at: https://www.icomos.org/en/charters-and-texts/179-articles-en-francais/ressources/charters-and-standards/170-european-charter-of-the-architectural-heritage [Accessed 11 February 2021].

Council of Europe. (2000). *European landscape convention (Florence Convention) European Treaty Series 176*. Strasbourg : Council of Europe. [Online]. Available at: https://www.coe.int/en/web/conventions/full-list [Accessed 10 December 2020].

Council of Europe. (2005). *Framework convention on the value of cultural heritage for society*. [Online]. Available at: https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/199 [Accessed 4 February 2021].

Countryside Agency. (2002). *Landscape character assessment. Guidance for England and Scottland*. The Countryside Agency. [Online]. Available at: http://publications.naturalengland.org.uk/publication/4670824246149120 [Accessed 27 January 2021].

Cox, K. R. (2014). *Making human geography*. New York, London: The Guilford Press.

Craig, W. J., Harris, T. M. and Weiner, D. (2002). *Community participation and geographic information systems*. London: Taylor & Francis.

Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. New Haven: Yale University Press.

Crawford, O. (1953). *Archaeology in the field*. London: Phoenix House.

Cresswell, T. (2015). *Place: an introduction*. 2nd ed. Chichester; Oxford: Wiley Blackwell.

Creswell, J. W. author. (2017). *Research design: Qualitative, quantitative , and mIxed methods approaches*. London: SAGE Publications.

Crouch, D. (1996). Making sense of our place: a critical review of parish maps. In: Clifford, S. and King, A. (Eds). *from place to PLACE - maps and parish maps*. London : Common Ground. pp.53–65. [Online]. Available at: https://www.commonground.org.uk/shop/from-place-to-place/ [Accessed 6 December 2020].

Currie, M. and Correa, M. M. (2022). Tangibles, intangibles and other tensions in the culture and communities mapping project. *Cultural Trends*, 31 (1), pp.88–106. [Online]. Available at: doi:10.1080/09548963.2021.1910491.

Currie, M. and Miranda Correa, M. (2021). *The culture and communities mapping project*. Cham: Springer International Publishing. [Online]. Available at: doi:10.1007/978-3-030-88651-6 [Accessed 26 November 2023].

Dabaut, N. (2021). *Landscape character and public perception: a participatory historic landscape characterisation approach in Northumberland*. Newcastle University.

Dalglish, C. (2018). *Community empowerment and landscape*. Glasgow: Inherit and Community Land Scotland.

Dalglish, C. and Leslie, A. (2016). A question of what matters: landscape characterisation as a process of situated, problem-orientated public discourse. *Landscape Research*, 41 (2), pp.212–226. [Online]. Available at: doi:10.1080/01426397.2015.1135319.

Darvill, T. (1999). The historic environment, historic landscapes, and space-time-action modelss in landscape archaeology. In: Ucko, P. J. and Layton, R. (Eds). *The archaeology and anthropology of landscape: shaping your landscape*. One world archaeology. London, etc.: Routledge. pp.104–118.

Dobson, S. and Selman, P. (2012). Applying historic landscape characterization in spatial planning: from remnants to remanence. *Planning Practice and Research*, 27 (4), pp.459–474. [Online]. Available at: doi:10.1080/02697459.2012.680268.

Drus, Z. and Khalid, H. (2019). Sentiment analysis in social media and its application: systematic literature review. *Procedia Computer Science*, 161, pp.707–714. [Online]. Available at: doi:10.1016/j.procs.2019.11.174.

Eagles, P. (2022). Artificial intelligence for data enhancement, linking and exploration. *https://unpathdwaters.org.uk*. [Online]. Available at: https://unpathdwaters.org.uk/artificial-intelligence-for-data-enhancement-linking-and-exploration/ [Accessed 2 November 2023].

Earley-Spadoni, T. (2017). Spatial history, deep mapping and digital storytelling: archaeology's future imagined through an engagement with the digital humanities. *Journal of Archaeological Science*, 84, pp.95–102. [Online]. Available at: doi:10.1016/j.jas.2017.05.003.

Edelson, L. and McCoy, D. (2021). Facebook is obstructing our work on disinformation. Other researchers could be next. [Online]. Available at: https://www.theguardian.com/technology/2021/aug/14/facebook-research-disinformation-politics [Accessed 11 February 2022].

Emerick, K. (2014). *Conserving and managing ancient monuments: heritage, democracy, and inclusion*. In: Heritage Matters. Boydell & Brewer. [Online]. Available at:

https://www.cambridge.org/core/books/conserving-and-managing-ancient-monuments/99AEFF7CEDEA0F8CB53D131E54634F8D [Accessed 29 June 2023].

Emerick, K. (2016). The language changes but practice stays the same. Does the same have to be true for community conservation? In: Chitty, G. (Ed). *Heritage, conservation and communities: engagement, participation and capacity building*. Heritage, culture, and identity. London: Routledge. pp.65–77.

English Heritage. (1997). *Sustaining the historic environment: new perspectives on the future.* London: English Heritage.

English Heritage. (2000). *Power of place: the future of the historic environment*. London: English Heritage.

English Heritage. (2008). *Conservation principles, policies and guidance*. Swindon: English Heritage.

English Heritage. (2011a). *Knowing your place*. [Online]. Available at: https://www.stratford.gov.uk/doc/173665/name/English%20Heritage%20Knowing%20your%20Place.pdf/ [Accessed 13 February 2021].

English Heritage. (2011b). *Our places*. [Online]. Available at: https://www.english-heritage.org.uk/about-us/our-places/ [Accessed 13 February 2021].

Exegesis and Locus Consulting. (2017). *National HLC stage 5: final report*. Natural England.

Fagerholm, N. et al. (2022). Analysis of pandemic outdoor recreation and green infrastructure in Nordic cities to enhance urban resilience. *npj Urban Sustainability*, 2 (1), pp.1–14. [Online]. Available at: doi:10.1038/s42949-022-00068-8.

Fairclough, G. (2001). Cultural landscape, computers and characterisation: GIS-based historic landscape characterisation as a tool for archaeological resource management in England. In: Burenhult, G. and Arvidson, J. (Eds). *Computer applications and quantitative methods in archaeology. Proceedings of the 29th Conference, Gotland, April 2001 (BAR International Series 1016).* 2001. Oxford : Archaeopress. pp.277–294.

Fairclough, G. (2002). Europe's landscape: archaeology, sustainability and agriculture. In: Fairclough, G. J., Rippon, S. and Bull, D. (Eds). *Europe's Cultural Landscape: archaeologists and the management of change*. Brussels: Europae Archaeologiae Consilium. pp.1–12.

Fairclough, G. (2007). The contemporary and future landscape: change and creation in the later twentieth century. In: McAtackney, L., Palus, M. M. and Piccini, A. (Eds). *Contemporary and historical archaeology in theory: papers from the 2003 and 2004 CHAT conferences*. BAR international series 1677. Oxford : Archaeopress.

Fairclough, G. J. et al. (1999). *Yesterday's world, tomorrow's landscape : the English Heritage historic landscape project 1992-94*. London : English Heritage.

Fairclough, G. J. and Barnatt, J. (1999). *Historic landscape characterisation: 'the state of the art' : papers from a seminar held at Society of Antiquaries, Burlington House, London, 11 December 1998*. London: English Heritage.

Farrow, J. (2021). *May 2021 - Cold and wet - saved only by the bank holiday weekend*. [Online]. Available at: https://www.netweather.tv/weather-forecasts/news/10895-may-2021---cold-and-wet---saved-only-by-the-bank-holiday-weekend [Accessed 11 February 2022].

Feld, S. and Basso, K. (1996). *Senses of place*, School of American Research advanced seminar series. Santa Fe, N.M.: [Seattle] : School of American Research Press ; Distributed by the University of Washington Press.

Finch, J. (2007a). "What more were the pastures of Leicester to me?" Hunting, landscape character, and the politics of place. *International Journal of Cultural Property*, 14 (3), pp.361–383. [Online]. Available at: doi:10.1017/S0940739107070233.

Finch, J. (2007b). 'Wider famed countries': historic landscape characterisation in the Midland shires. *Landscapes*, 8 (2), pp.50–63.

Fiorucci, M. et al. (2020). Machine learning for cultural heritage: A survey. *Pattern Recognition Letters*, 133, pp.102–108. [Online]. Available at: doi:10.1016/j.patrec.2020.02.017.

Flybjerg, B. (2011). Case study. In: Denzin, N. K. and Lincoln, Y. S. (Eds). *The Sage handbook of qualitative research*. London: Thousand Oaks, Sage. pp.301–316.

Franzosi, R., Dong, W. and Dong, Y. (2022). Qualitative and quantitative research in the humanities and social sciences: how natural language processing (NLP) can help. *Quality & Quantity*, 56, pp.1–31. [Online]. Available at: doi:10.1007/s11135-021-01235-2.

Fredengren, C. (2015). NATURE:CULTURES Heritage, sustainability and feminist posthumanism. *Current Swedish Archaeology*, 23, pp.109–130.

Fredheim, L. H. and Khalaf, M. (2016). The significance of values: heritage value typologies re-examined. *International Journal of Heritage Studies*, 22 (6), pp.466–481. [Online]. Available at: doi:10.1080/13527258.2016.1171247.

Gaffney, V. and Dingwall, L. (Eds). (2007). *Heritage management at fort hood, Texas: Experiments in historic landscape characterisation*. Archaeopress.

Gajadhar, J. and Green, J. (2005). *The importance of nonverbal elements in online chat*. p.2.

Gard'ner, J. M. (2004). Heritage protection and social inclusion: a case study from the Bangladeshi community of East London. *International Journal of Heritage Studies*, 10 (1), pp.75–92. [Online]. Available at: doi:10.1080/1352725032000194259.

Gatersleben, B. et al. (2020). Why are places so special? Uncovering how our brain reacts to meaningful places. *Landscape and Urban Planning*, 197, p.103758. [Online]. Available at: doi:10.1016/j.landurbplan.2020.103758.

Gattiglia, G. (2022). A postphenomenological perspective on digital and algorithmic archaeology. In: *Archeologia e Calcolatori*. 33(2). All'Insegna del Giglio. pp.319–334.

Geertz, C. (1973). *The interpretation of cultures [electronic resource]: selected essays*, ACLS Humanities E-Book. New York : Basic Books. [Online]. Available at: http://hdl.handle.net/2027/heb.01005 [Accessed 19 January 2021].

Gibney, E. (2019). Privacy hurdles thwart Facebook democracy research. *Nature*, 574 (7777), pp.158–159. [Online]. Available at: doi:10.1038/d41586-019-02966-x.

Ginzarly, M. and Srour, F. (2021). Cultural heritage through the lens of COVID-19. *Poetics*, p.101622. [Online]. Available at: doi:10.1016/j.poetic.2021.101622.

Goerz, G. and Scholz, M. (2010). Adaptation of NLP Techniques to Cultural Heritage Research and Documentation. *Journal of Computing and Information Technology*, 18 (4), p.317.

Golledge, R. G. (2006). Philosophical bases of behavioural research in geography[ebook]. In: Aitken, S. C. and Valentine, G. (Eds). *Approaches to Human Geography*. London: SAGE Publications. p.7585.

Golzar, J. and Tajik, O. (2022). Convenience sampling. *International Journal of Education and Language Studies*, 1 (2), pp.72–77.

Graham, B., Ashworth, G. J. and Turnbridge, J. E. (2000). *A Geography of Heritage*. New York: Oxford Universtiy Press.

Graham, H., Mason, R. and Newman, A. (2009). *Literature review: historic environment, sense of place and social capital*. Newcastle : ICCHS, Newcastle University., p.44. [Online]. Available at: https://historicengland.org.uk/content/heritage-counts/pub/sense_of_place_lit_review_web1-pdf/.

Greider, T. and Garkovich, L. (1994). Landscapes: the social construction of nature and the environment. *Rural Sociology*, 59, pp.1–24. [Online]. Available at: doi:10.1111/j.1549-0831.1994.tb00519.x.

Grenville, J. (2007). Conservation as psychology: ontological security and the built environment. *International Journal of Heritage Studies*, 13 (6), pp.447–461. [Online]. Available at: doi:10.1080/13527250701570614.

Grove-White, R. (1996). Parish maps: local knowledge and the reconstruction of democracy. In: Clifford, S. and King, A. (Eds). *from place to PLACE - maps and parish maps*. London: Common Ground. pp.9–14. [Online]. Available at: https://www.commonground.org.uk/shop/from-place-to-place/ [Accessed 6 December 2020].

Gutherz, G. et al. (2023). Translating Akkadian to English with neural machine translation. *PNAS Nexus*, 2 (5), p.pgad096. [Online]. Available at: doi:10.1093/pnasnexus/pgad096.

Gutowski, P. and Kłos-Adamkiewicz, Z. (2020). Development of e-service virtual museum tours in Poland during the SARS-CoV-2 pandemic. *Procedia Computer Science*, 176, pp.2375–2383. [Online]. Available at: doi:10.1016/j.procs.2020.09.303.

Harari, Y. N. (2017). The rise of the useless class. *ideas.ted.com*. [Online]. Available at: https://ideas.ted.com/the-rise-of-the-useless-class/ [Accessed 17 May 2023].

Harris, T. M. (2015). Deep geography - deep mapping. Spatial storytelling and a sense of place. In: Bodenhamer, D. J. et al. (Eds). *Deep maps and spatial narratives*. Bloomington: Indiana University Press. pp.28–53. [Online]. Available at: http://ebookcentral.proquest.com/lib/york-ebooks/detail.action?docID=4697517 [Accessed 8 January 2021].

Harrison, R. (2010). Heritage as social action. In: West, S. (Ed). *Understanding heritage in practice*. Understanding global heritage. Manchester: University Press. pp.240–276.

Harrison, R. (2011). 'Counter-mapping' heritage, communities and places in Australia and the UK. In: Schofield, J. and Szymanski, R. (Eds). *Local heritage, global context: cultural perspectives on sense of place. Heritage, culture and identity*. Farnham: Ashgate. pp.79–98. [Online]. Available at: http://www.ashgate.com/default.aspx?page=637&calcTitle=1&title_id=9327&edition_id=12403 [Accessed 29 October 2020].

Harrison, R. (2015). Beyond "natural" and "cultural" heritage: toward an ontological politics of heritage in the age of anthropocene. *Heritage & Society*, 8 (1), pp.24–42. [Online]. Available at: doi:10.1179/2159032X15Z.00000000036.

Harrison, R. (2020). *Heritage futures: comparative approaches to natural and cultural heritage practices*. UCL Press. [Online]. Available at: https://play.google.com/books/reader?id=LVv5DwAAQBAJ&hl=en&pg=GBS.PA3.

Hausmann, A. et al. (2020). Understanding sentiment of national park visitors from social media data. Puig De La Bellacasa, M. (Ed). *People and Nature*, 2 (3), pp.750–760. [Online]. Available at: doi:10.1002/pan3.10130.

Hayden, D. (1995). *The power of place*. Cambridge, MA : The MIT Press. [Online]. Available at: https://mitpress.mit.edu/books/power-place [Accessed 26 December 2020].

Hedblom, M. et al. (2020). Landscape perception: linking physical monitoring data to perceived landscape properties. *Landscape Research*, 45 (2), pp.179–192. [Online]. Available at: doi:10.1080/01426397.2019.1611751.

Hegelich, S. (2020). Facebook needs to share more with researchers. *Nature*, 579 (7800), pp.473–473. [Online]. Available at: doi:10.1038/d41586-020-00828-5.

Hern, A. (2015). Don’t know the difference between emoji and emoticons? Let me explain. [Online]. 6 February. Available at: https://www.theguardian.com/technology/2015/feb/06/difference-between-emoji-and-emoticons-explained [Accessed 11 February 2022].

Herring, P. C. (2009). Framing perceptions of the historic landscape: historic landscape characterisation (HLC) and historic land-use assessment (HLA). *Scottish Geographical Journal*, 125 (1), pp.61–77. [Online]. Available at: doi:10.1080/14702540902873907.

Historic England. (2021). *Historic landscape characterisation*. [Online]. Available at: https://historicengland.org.uk/research/methods/characterisation/historic-landscape-characterisation/.

Historic England. (2022a). *Download listing data - GIS shapefiles*. [Online]. Available at: http://historicengland.org.uk/listing/the-list/data-downloads/ [Accessed 12 February 2022].

Historic England. (2022b). *Survey of COVID-19 effects on the heritage sector*. [Online]. Available at: http://historicengland.org.uk/coronavirus/heritage-sector/survey/ [Accessed 13 December 2022].

Historic England. (2023). *Historic England corporate plan 2023-26*. [Online]. Available at: https://historicengland.org.uk/images-books/publications/he-corp-plan-2023-26/.

Hølleland, H. and Skrede, J. (2019). What’s wrong with heritage experts? An interdisciplinary discussion of experts and expertise in heritage studies. *International Journal of Heritage Studies*, 25 (8), pp.825–836. [Online]. Available at: doi:10.1080/13527258.2018.1552613.

Hooley, D. (2014). *Historic seascape characterisation South West Peninsula*. English Heritage and Cornwall Council.

Hoskins, W. G. (1955). *The making of the English landscape*. London: Hodder & Stoughton.

Hoskins, W. G. (1977). *The making of the English landscape.* London, etc.: Hodder & Stoughton.

Huggett, J. (2021). Algorithmic agency and autonomy in archaeological practice. *Open Archaeology*, 7, pp.417–434. [Online]. Available at: doi:10.1515/opar-2020-0136.

Hunziker, M., Buchecker, M. and Hartig, T. (2007). *Space and place – two aspects of the human-landscape relationship*. In: pp.47–62. [Online]. Available at: doi:10.1007/978-1-4020-4436-6_5.

Hutto, C. and Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8 (1), pp.216–225.

Hutto, C. J. (2022). *cjhutto/vaderSentiment*. Python. [Online]. Available at: https://github.com/cjhutto/vaderSentiment [Accessed 14 January 2022].

ICOMOS. (1964). *International charter for the conservation and restoration of monuments and sites.* Venice : ICOMOS. [Online]. Available at: https://www.icomos.org/charters/venice_e.pdf.

ICOMOS. (1975). *The declaration of Amsterdam - 1975 - International Council on Monuments and Sites*. In: 1975. Amsterdam : ICOMOS - International Council on Monuments and Sites. [Online]. Available at: https://www.icomos.org/en/and/169-the-declaration-of-amsterdam [Accessed 16 November 2022].

ICOMOS. (1979). *Burra-Charter_1979*. Australia : ICOMOS. [Online]. Available at: https://australia.icomos.org/wp-content/uploads/Burra-Charter_1979.pdf [Accessed 14 December 2022].

ICOMOS. (2013). *The Burra Charter: the Australia ICOMOS charter for places of cultural significance 2013*. [Online]. Available at: http://australia.icomos.org/wp-content/uploads/The-Burra-Charter-2013-Adopted-31.10.2013.pdf [Accessed 16 October 2020].

ICOMOS. (2021a). *Declaration of Dresden on the 'Reconstruction of monuments destroyed by war' (1982) - International Council on Monuments and Sites*. [Online]. Available at: https://www.icomos.org/en/charters-and-texts/179-articles-en-francais/ressources/charters-and-standards/184-the-declaration-of-dresden [Accessed 11 February 2021].

ICOMOS. (2021b). *The Athens charter for the restoration of historic monuments - 1931*. [Online]. Available at: https://www.icomos.org/en/167-the-athens-charter-for-the-restoration-of-historic-monuments [Accessed 10 February 2021].

Iglesias, C. and Moreno, A. (Eds). (2020). *Sentiment analysis for social media*. MDPI. [Online]. Available at: https://www.mdpi.com/books/pdfview/book/2154 [Accessed 28 January 2022].

Ihde, D. (2009). *Postphenomenology and technoscience: the Peking University lectures*, SUNY series in the philosophy of the social sciences. New York: Suny Press. [Online]. Available at: https://sunypress.edu/Books/P/Postphenomenology-and-Technoscience [Accessed 31 October 2023].

Ingold, T. (1993). The temporality of the landscape. *World Archaeology*, 25 (2), pp.152–174.

Institute for Government Analysis. (2021). *timeline-lockdown*. [Online]. Available at: https://www.instituteforgovernment.org.uk/sites/default/files/timeline-lockdown-web.pdf [Accessed 25 February 2022].

Ireland, T., Brown, S. and Schofield, J. (2020). Situating (in)significance. *International Journal of Heritage Studies*, 26 (9), pp.826–844. [Online]. Available at: doi:10.1080/13527258.2020.1755882.

Ireland, T. and Schofield, J. (2015). *The ethics of cultural heritage*, Ethical Archaeologies: The Politics of Social Justice. Springer.

Jacobs, M. (2006). *The production of mindscapes: a comprehensive theory of landscape experience. Thesis*. Wageningen: University of Wageningen.

Jahn, M. and Buchholz, S. (2010). Space in narrative. In: *Routledge encyclopedia of narrative theory*. Taylor and Francis. [Online]. Available at: doi:10.4324/9780203932896.

Jeffries, J. (2012). Pattern, patterning. In: Lury, C. and Wakeford, N. (Eds). *Inventive methods: the happening of the social*. Culture, economy and the social. New York, London : Routledge. pp.126–135. [Online]. Available at: https://ebookcentral.proquest.com/lib/york-ebooks/detail.action?docID=981968.

Johnson, R. L. et al. (2022). *The ghost in the machine has an american accent: value conflict in GPT-3*. arXiv. [Online]. Available at: http://arxiv.org/abs/2203.07785 [Accessed 26 April 2023].

Johnston, C. (1992). *What is social value? A Discussion Paper*. Canberra : Australian Government Publishing Service. [Online]. Available at: https://www.academia.edu/1098089/What_is_Social_Value_A_Discussion_Paper [Accessed 23 October 2020].

Johnston, C. (2017). Recognising connection: social significance and heritage practice. *Córima*, 2 (2). [Online]. Available at: https://www.academia.edu/36157277/Recognising_connection_social_significance_and_heritage_practice [Accessed 24 October 2020].

Johnston, C. (2023). Social value. Identifying, documenting, and assessing community connections. In: Brown, S. and Goetcheus, C. (Eds). *Routledge Handbook of Cultural Landscape Practice*. London: Routledge. pp.245–259.

Jokilehto, J. (2016). Engaging conservation: community, place and capacity building. In: Chitty, G. (Ed). *Heritage, conservation and communities: engagement, participation and capacity building*. Heritage, culture, and identity. London: Routledge. pp.17–33.

Jones, N. and McGinlay, J. (2020). *The impact of COVID-19 restrictions on local communities of Peak District National Park and management options during the pandemic*. Cambridge: University of Cambridge., p.20.

Jones, S. (2017). Wrestling with the social value of heritage: problems, dilemmas and opportunities. *Journal of Community Archaeology & Heritage*, 4 (1), Routledge., pp.21–37. [Online]. Available at: doi:10.1080/20518196.2016.1193996.

Jones, S. et al. (2024). Assessing the dynamic social values of the ‘deep city’: An integrated methodology combining online and offline approaches. *Progress in Planning*, p.100852. [Online]. Available at: doi:10.1016/j.progress.2024.100852.

Jones, S. and Leech, S. (2015). *Valuing the historic environment: a critical review of existing approaches to social value. AHRC Cultural Value Report*. Manchester: University of Manchester. [Online]. Available at: https://www.escholar.manchester.ac.uk/uk-ac-man-scw:281849 [Accessed 26 October 2020].

Jones, T., Doane, W. and Attbom, M. (2021). *textmineR: functions for text mining and topic modeling. Version 3.0.5*. [Online]. Available at: https://CRAN.R-project.org/package=textmineR [Accessed 19 August 2022].

Jones, T. W. (2021). *Topic modeling*. [Online]. Available at: https://cran.r-project.org/web/packages/textmineR/vignettes/c_topic_modeling.html [Accessed 31 July 2022].

Kansteiner, W. (2022). Digital doping for historians: can history, memory, and historical theory be rendered artificially InTELligent? *History and Theory*, 61 (4), pp.119–133. [Online]. Available at: doi:10.1111/hith.12282.

Kaplan, R. (1984). Impact of urban nature: A theoretical analysis. *Urban Ecology*, 8 (3), pp.189–197. [Online]. Available at: doi:10.1016/0304-4009(84)90034-2.

Kaufman, N. (2013). Putting intangible heritage in its place(s): proposals for policy and practice. *IJIH - International Journal of Intangible Heritage*, 8, pp.20–36.

Kiddey, R. (2014). Punks and drunks: counter-mapping homelessness in Bristol and York. In: Schofield, J. (Ed). *Who needs experts? Counter-mapping cultural heritage*. Abingdon: Routledge. pp.165–180.

Koblet, O. and Purves, R. S. (2020). From online texts to landscape character assessment: collecting and analysing first-person landscape perception computationally. *Landscape and Urban Planning*, 197, p.103757. [Online]. Available at: doi:10.1016/j.landurbplan.2020.103757.

Köpf, A. et al. (2023). *OpenAssistant conversations -- democratizing large language model alignment*. arXiv. [Online]. Available at: doi:10.48550/arXiv.2304.07327 [Accessed 10 May 2023].

Kwan, M.-P. and Ding, G. (2008). Geo-narrative: extending geographic information systems for narrative analysis in qualitative and mixed-method research∗. *The Professional Geographer*, 60 (4), pp.443–465. [Online]. Available at: doi:10.1080/00330120802211752.

Latour, B. (1993). *We have never been modern*. Cambridge, Mass: Harvard University Press.

Lau, J. H., Newman, D. and Baldwin, T. (2014). Machine reading tea leaves: automatically evaluating topic coherence and topic model quality. In: *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*. April 2014. Gothenburg, Sweden : Association for Computational Linguistics. pp.530–539. [Online]. Available at: doi:10.3115/v1/E14-1056 [Accessed 14 November 2022].

Leeson, W. et al. (2019). Natural language processing (NLP) in qualitative public health research: a proof of concept study. *International Journal of Qualitative Methods*, 18, p.1609406919887021. [Online]. Available at: doi:10.1177/1609406919887021.

Lewicka, M. (2011a). On the varieties of people's relationships with places: Hummon's typology revisited. *Environment and Behavior*, 43, pp.676–709. [Online]. Available at: doi:10.1177/0013916510364917.

Lewicka, M. (2011b). Place attachment: how far have we come in the last 40 years? *Journal of Environmental Psychology*, 31 (3), pp.207–230. [Online]. Available at: doi:10.1016/j.jenvp.2010.10.001.

Linebaugh, K. and Knutson, R. *The clash between Facebook and independent researchers - The Journal. - WSJ Podcasts*. [Online]. Available at: https://www.wsj.com/podcasts/the-journal/the-clash-between-facebook-and-independent-researchers/8fda97fc-203d-4632-bc86-f81c0cbe5faf [Accessed 11 February 2022].

Low, S. M. (1987). A cultural landscapes mandate for action. *CRM Bulletin*, 10 (1), pp.30–33.

Low, S. M. (2002). Anthropological-ethnographic methods for the assessment of cultural values in heritage conservation. In: Torre de la, M. (Ed). *Assessing the values of cultural heritage: Research report*. Los Angeles: The Getty Conservation Institute. pp.31–49.

Lynch, K. (1960). *The image of the city*. Cambridge: The MIT Press & Harvard University Press.

Lynch, K. (1972). *What time is this place?* Cambridge, Mass.: MIT Press.

Madgin, R. (2021). Emoji as method. In: Madgin, R. and Lech, J. (Eds). *People-centred methodologies for heritage conservation*. London: Routledge. pp.80–94.

Madgin, R. and Lesh, J. (2021). *People-centred methodologies for heritage conservation | Taylor & Francis Group*. [Online]. Available at: https://www.taylorfrancis.com/https://www.taylorfrancis.com/books/edit/10.4324/9780429345807/people-centred-methodologies-heritage-conservation-rebecca-madgin-james-lesh.

Madgin, R. and Robson, E. (2023). *Developing a people-centred, place-led approach: the value of the arts and humanities*. University of Glasgow.

Maguire, B. D. (2017). *Modeling place attachment using GIS*. Vancouver: University of British Columbia. [Online]. Available at: https://open.library.ubc.ca/media/download/pdf/24/1.0348807/3 [Accessed 13 August 2021].

Mahmud, J., Nichols, J. and Drews, C. (2014). Home location identification of Twitter users. *ACM Transactions on Intelligent Systems and Technology*, 5 (3), pp.1–21. [Online]. Available at: doi:10.1145/2528548.

Malde, R. (2020). *A short introduction to VADER*. [Online]. Available at: https://towardsdatascience.com/an-short-introduction-to-vader-3f3860208d53 [Accessed 30 August 2021].

Manzo, L. editor of compilation and Devine-Wright, P. (2014). *Place attachment [electronic resource]: advances in theory, methods and applications*. London: Routledge.

Manzo, L. and Perkins, D. (2006). Finding common ground: the importance of place attachment to community participation and planning. *Journal of Planning Literature*, 20, pp.335–350. [Online]. Available at: doi:10.1177/0885412205286160.

Mason, R. (1999). *Economics and heritage conservation: a meeting organized by the GCI, December 1998*. Los Angeles : Getty Conservation Institute. [Online]. Available at: https://www.getty.edu/conservation/publications_resources/pdf_publications/economics_and_heritage.html [Accessed 22 January 2021].

Mason, R. (2002). Assessing the values in conservation planning: methodological issues and choices. In: De La Torre, M. (Ed). *Assessing the values of cultural heritage: Research Report*. Los Angeles: The Getty Conservation Institute. pp.5–30.

Matrone, F. et al. (2020). Comparing machine and deep learning methods for large 3D heritage semantic segmentation. *ISPRS International Journal of Geo-Information*, 9 (9), Multidisciplinary Digital Publishing Institute., p.535. [Online]. Available at: doi:10.3390/ijgi9090535.

Mayhew, S. (2009). Landscape. In: *A Dictionary of Geography*. Oxford University Press. [Online]. Available at: https://www.oxfordreference.com/view/10.1093/acref/9780199231805.001.0001/acref-9780199231805-e-1796 [Accessed 14 February 2021].

McKeithen, W. (2015). By-roads and hidden treasures: mapping cultural assets in regional Australia. *Journal of Cultural Geography*, 32, pp.1–3. [Online]. Available at: doi:10.1080/08873631.2015.1069501.

Meinig, D. W. (1979). The behholding eye: Ten versions of the same scene. In: Jackson, J. B. and Meinig, D. W. (Eds). *The interpretation of ordinary landscapes: geographical essays*. New York: Oxford University Press. pp.33–48.

Meinig, D. W. and Jackson, J. B. (Eds). (1979). *The interpretation of ordinary landscapes: geographical essays*. New York : Oxford University Press.

Messina, C. (2022). *Chris Messina*. [Online]. Available at: https://chrismessina.me [Accessed 28 January 2022].

Meta. (2021). The Facebook company Is now Meta. *Meta*. [Online]. Available at: https://about.fb.com/news/2021/10/facebook-company-is-now-meta/ [Accessed 11 February 2022].

Meta. (2022). *Academic resources - Meta research | Meta Research*. [Online]. Available at: https://research.facebook.com/data/ [Accessed 11 February 2022].

MetOffice. (2019). *Weather this bank holiday weekend?* [Online]. Available at: https://www.metoffice.gov.uk/about-us/press-office/news/weather-and-climate/2019/late-may-bank-holiday-2019 [Accessed 11 February 2022].

MetOffice. (2020). *May 2020 becomes the sunniest calendar month on record*. [Online]. Available at: https://www.metoffice.gov.uk/about-us/press-office/news/weather-and-climate/2020/2020-spring-and-may-stats [Accessed 11 February 2022].

Milgram, S. and Jodelet, D. (1992). Psychological maps of Paris. In: Milgram, S. (Ed). *The individual in a social world. Essays and Experiments*. Second. McGraw-Hill. pp.88–113.

Milligan, M. J. (1998). Interactional past and potential: the social construction of place attachment. *Symbolic Interaction*, 21 (1), [Wiley, Society for the Study of Symbolic Interaction]., pp.1–33. [Online]. Available at: doi:10.1525/si.1998.21.1.1.

Modesto, G. and Waterton, E. (2020). The elite and the everyday in the Australian heritage field. In: Bennett, T. et al. (Eds). *Fields, capitals, habitus: Australian culture, inequalities and sociald*. 1st ed. Routledge. [Online]. [Accessed 15 April 2023].

Monmonier, M. (1996). *How to lie with maps*. Second. Chicago: University of Chicago Press.

Nardi, S. D. (2014). Senses of place, senses of the past: making experiential maps as part of community heritage fieldwork. *Journal of Community Archaeology & Heritage*, 1 (1), pp.5–22. [Online]. Available at: doi:10.1179/2051819613Z.0000000001.

National Museum Liverpool. (2021). *Covid-19 display*. [Online]. Available at: https://www.liverpoolmuseums.org.uk/stories/covid-19-display [Accessed 22 August 2021].

National Trust. (2017). *places-that-make-us-research-report.pdf*. Places that make us, Swindon : National Trust. [Online]. Available at: https://nt.global.ssl.fastly.net/documents/places-that-make-us-research-report.pdf [Accessed 3 May 2021].

Natural England. (2015). *Econets, landscape & people: Integrating people's values and cultural ecosystem services into the design of ecological networks and other landscape change proposals - NECR180*. [Online]. Available at: http://publications.naturalengland.org.uk/publication/6172716216352768 [Accessed 23 November 2022].

Natural England. (2020). *National historic landscape characterisation 250m Grid (England)*. [Online]. Available at: https://data.gov.uk/dataset/8b8c5df3-d7e3-484c-89d8-c7b819205002/national-historic-landscape-characterisation-250m-grid-england [Accessed 14 February 2021].

Natural Resources Wales. (2016a). *LANDMAP guidance note 4: LANDMAP and the cultural landscape*. [Online]. Available at: https://cdn.cyfoethnaturiol.cymru/media/677880/landmap-guidance-note-4-landmap-and-the-cultural-landscape-2016.pdf?mode=pad&rnd=131471905520000000 [Accessed 19 February 2021].

Natural Resources Wales. (2016b). *LANDMAP methodology guidance note for Wales cultural landscapes*. [Online]. Available at: https://naturalresources.wales/media/676228/cultural-landscape-landmap-methodology-2016.pdf [Accessed 10 February 2021].

NatureScot. (2020). *Talking about our place toolkit*. [Online]. Available at: https://www.nature.scot/enjoying-outdoors/communities-and-landscape/talking-about-our-place-toolkit [Accessed 22 August 2021].

Neri, F. et al. (2012). *Sentiment analysis on social media*. In: 28 August 2012. [Online]. Available at: doi:10.1109/ASONAM.2012.164.

Nold, C. (Ed). (2009). *Emotional cartography: technologies of the self*. Creative Commons.

Novak, P. K. et al. (2015). Sentiment of emojis. *PLOS ONE*, 10 (12), Public Library of Science., p.e0144296. [Online]. Available at: doi:10.1371/journal.pone.0144296.

Novaković, P. (2018). *Impact of the large-scale excavations in the Slovene preventive archaeology*.

Odacioglu, E. C. and Zhang, L. (2022). Text mining for rendering theory: integrating topic modeling to grounded theory. *SSRN Electronic Journal*. [Online]. Available at: doi:10.2139/ssrn.4141372 [Accessed 24 September 2022].

Ordnance Survey. (n.d.). *Free OS OpenData map downloads*. [Online]. Available at: https://osdatahub.os.uk/downloads/open [Accessed 17 January 2022].

Oxford Languages. (2022a). *Oxford word of the year 2015 | Oxford Languages*. [Online]. Available at: https://languages.oup.com/word-of-the-year/2015/ [Accessed 26 January 2022].

Oxford Languages. (2022b). *Oxford word of the year 2020 | Oxford Languages*. [Online]. Available at: https://languages.oup.com/word-of-the-year/2020/ [Accessed 26 January 2022].

Panko, B. (n.d.). *A decade ago, the hashtag reshaped the internet*. [Online]. Available at: https://www.smithsonianmag.com/smart-news/decade-ago-hashtag-reshaped-internet-180964605/ [Accessed 28 January 2022].

PDNPA. (2009). *Peak District landscape strategy 2009 - Introduction*. [Accessed 1 January 2018].

PDNPA. (2021). *Neighbourhood areas*. [Online]. Available at: https://www.peakdistrict.gov.uk/planning/neighbourhood-planning/neighbourhood-areas [Accessed 6 February 2021].

PDNPA. (n.d.). *History of our national park: Peak District National Park*. [Online]. Available at: http://www.peakdistrict.gov.uk/learning-about/about-the-national-park/our-history [Accessed 3 January 2022a].

PDNPA. (n.d.). *Peak District facts*. [Online]. Available at: https://www.peakdistrict.gov.uk/learning-about/news/70-years-of-the-peak-district-national-park/peak-district-facts [Accessed 13 December 2022b].

Pearson, M. (1995). *Looking after heritage places: the basics of heritage planning for managers, landowners and administrators*. Carlton, Vic.: Melbourne University Press.

Pearson, M. and Shanks, M. (2001). *Theatre/archaeology*. London: Routledge.

Pendlebury, J. and Gibson, L. (2016). Introduction: valuing historic environments. In: *Valuing historic environments*. 2016 : Routledge. pp.1–18. [Online]. Available at: doi:10.4324/9781315548449 [Accessed 28 September 2020].

Perkins, C. (2007). Community mapping. *Cartographic Journal, The*, 44, pp.127–137. [Online]. Available at: doi:10.1179/000870407X213440.

Petrova, S., Čihař, M. and Bouzarovski, S. (2011). Local nuances in the perception of nature protection and place attachment: a tale of two parks. *Area*, 43 (3), pp.327–335. [Online]. Available at: doi:10.1111/j.1475-4762.2011.00995.x.

Phillips, A. (2005). Landscape as a meeting ground: category v protected landscapes/seascapes and world heritage cultural landscapes. In: Brown, J., Mitchell, N. and Beresford, M. (Eds). *The Protected landscape approach: linking nature, culture and community*. Gland, Switzerland : IUCUN. pp.19–35. [Online]. Available at: https://portals.iucn.org/library/sites/library/files/documents/2005-006.pdf [Accessed 26 December 2020].

Piaget, J. (2013). *The construction of reality in the child*. London: Routledge. [Online]. Available at: doi:10.4324/9781315009650.

Pink, S. (2012). *Situating everyday life: practices and places*. London: SAGE Publications, Sage Publications.

Pink, S. author. (2016). *Digital ethnography: principles and practice*. Los Angeles; London: SAGE.

Pistilli, G. (2022). What lies behind AGI: ethical concerns related to LLMs. *Revue Ethique et Numérique*. [Online]. Available at: https://hal.science/hal-03607808 [Accessed 17 May 2023].

Polfliet, S. (2020). *Students place attachmen*t. Groningen, Netherlands: University of Groningen. [Online]. Available at: https://frw.studenttheses.ub.rug.nl/3209/1/Master%20thesis_Sara%20Polfliet_s3012964_students_place%20attachment.pdf [Accessed 29 August 2023].

Ponti, M. and Seredko, A. (2022). Human-machine-learning integration and task allocation in citizen science. *Humanities and Social Sciences Communications*, 9 (1), Palgrave., pp.1–15. [Online]. Available at: doi:10.1057/s41599-022-01049-z.

Primdahl, J. and Kristensen, L. S. (2016). Landscape strategy making and landscape characterisation—experiences from Danish experimental planning processes. *Landscape Research*, 41 (2), pp.227–238. [Online]. Available at: doi:10.1080/01426397.2015.1135322.

Proshansky, H. M., Fabian, A. K. and Kaminoff, R. (1983). Place-identity: Physical world socialization of the self. *Journal of Environmental Psychology*, 3 (1), pp.57–83. [Online]. Available at: doi:10.1016/S0272-4944(83)80021-8.

Pulido, C. et al. (2018). Social impact in social media: A new method to evaluate the social impact of research. *PLOS ONE*, 13, p.e0203117. [Online]. Available at: doi:10.1371/journal.pone.0203117.

Purves, R., Koblet, O. and Adams, B. (2022). *Analysing environmental narratives computationally*. In: Unlocking environmental narratives: towards understanding human environment interactions through computational text analysis, pp.43–84. [Online]. Available at: doi:10.5334/bcs.c.

Raji, I. D. et al. (2020). Saving face: investigating the ethical concerns of facial recognition auditing. In: *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. AIES '20. 7 February 2020. New York, NY, USA : Association for Computing Machinery. pp.145–151. [Online]. Available at: doi:10.1145/3375627.3375820 [Accessed 16 May 2023].

Ramkissoon, H., Weiler, B. and Smith, L. (2012). Place attachment and pro-environmental behaviour in national parks: The development of a conceptual framework. *Journal of Sustainable Tourism - J SUSTAIN TOUR*, 20, pp.257–276. [Online]. Available at: doi:10.1080/09669582.2011.602194.

Raymond, C. (2013). Review of place attachment: advances in theory, methods and applications (Edited by Lynne Manzo and Patrick Devine-Wright). *Estudios de Psicología*, 34 (3), pp.345–348. [Online]. Available at: doi:10.1174/021093913808349235.

RDocumentation. (n.d.). *findAssocs function - RDocumentation*. [Online]. Available at: https://www.rdocumentation.org/packages/tm/versions/0.7-8/topics/findAssocs [Accessed 28 January 2022].

Relph, E. (2021). *Placeness, place, placelessness*. [Online]. Available at: https://www.placeness.com/ [Accessed 22 August 2021].

Relph, E. C. (1976). *Place and placelessness*. London: Pion.

Resler, A. et al. (2021). A deep-learning model for predictive archaeology and archaeological community detection. *Humanities and Social Sciences Communications*, 8 (1), pp.1–10. [Online]. Available at: doi:10.1057/s41599-021-00970-z.

Riegl, A. (1903). *Der moderne Denkmalkultus. Sein Wesen und seine Entstehung.* [Online]. Available at: https://ia802708.us.archive.org/11/items/modernedenkmalk00denkgoog/modernedenkmalk00denkgoog.pdf [Accessed 9 December 2020].

Ritter, A. and Clark, S. (2011). Named entity recognition in tweets: An Experimental Study. In: *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*. 2011. Edinburgh, Scotland, UK : Association for Computational Linguistics. pp.1524–1534.

Robson, L. (2021). *Social value toolkit – guidance for heritage practitioners*. [Online]. Available at: https://socialvalue.stir.ac.uk/ [Accessed 6 October 2023].

Rogers, A., Castree, N. and Kitchin, R. (2013). Sauer, Carl O. In: *A dictionary ofhHuman geography*. Oxford University Press. [Online]. Available at: http://www.oxfordreference.com/view/10.1093/acref/9780199599868.001.0001/acref-9780199599868-e-1627 [Accessed 21 January 2021].

Rollero, C. and De Piccoli, N. (2010). Place attachment, identification and environment perception: An empirical study. *Journal of Environmental Psychology*, 30 (2), pp.198–205. [Online]. Available at: doi:10.1016/j.jenvp.2009.12.003.

Rudolff, B. (2006). *'Intangible' and 'tangible' heritage: A topology of culture in contexts of faith*. Mainz, Germany : Johannes-Gutenberg University. [Online]. Available at: http://icich.icomos.org/index.php/publication/intangible-and-tangible-heritage-a-topology-of-culture-in-contexts-of-faith/ [Accessed 19 October 2020].

Russell, S. and Norvig, P. (2016). *Artificial Intelligence: A Modern Approach, Global Edition*. 3rd edition. Boston: Pearson.

Sagan, C. (1990). The Prescott Courier - Google News Archive Search. *Parade Magazine*, Sept. 9, 1990. [Online]. [Accessed 16 August 2021].

Samaroudi, M., Echavarria, K. R. and Perry, L. (2020). Heritage in lockdown: digital provision of memory institutions in the UK and US of America during the COVID-19 pandemic. *Museum Management and Curatorship*, 35 (4), pp.337–361. [Online]. Available at: doi:10.1080/09647775.2020.1810483.

Samuels, A. and Mcgonical, J. (2020). Sentiment analysis on social media content. *arXiv:2007.02144 [cs]*. [Online]. Available at: http://arxiv.org/abs/2007.02144 [Accessed 28 January 2022].

Sassolini, E. and Cinini, A. (2010). Cultural heritage: Knowledge extraction from web documents. In: *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*. May 2010. Valletta, Malta : European Language Resources Association (ELRA). [Online]. Available at: http://www.lrec-conf.org/proceedings/lrec2010/pdf/415_Paper.pdf [Accessed 19 November 2022].

Sauer, C. O. (1996). The morphology of landscape. In: Agnew, J., Livingstone, D. N. and Rogers, A. (Eds). *Human Geography: an essential anthology*. Oxford : Blackwell. pp.297–315. [Online]. [Accessed 21 January 2021].

Scannell, L. and Gifford, R. (2010). The relations between natural and civic place attachment and pro-environmental behavior. *Journal of Environmental Psychology*, 30 (3), pp.289–297.

Schofield, J. (2007). Intimate engagements: arts, heritage and experience at the 'place-ballet'. *International Journal of Arts in Society*, 1 (5), pp.105–114.

Schofield, J. (2014). *Who needs experts? Counter-mapping cultural heritage*. Abingdon: Routledge.

Schofield, J. (2015a). Forget about 'heritage': place, ethics and the Faro convention. In: Schofield, J. and Ireland, T. (Eds). *The ethics of cultural heritage*. Ethical Archaeologies: The Politics of Social Justice. Springer. pp.197–209.

Schofield, J. (2015b). 'Thinkers and feelers': a psychological perspective on heritage and society. In: Waterton, E. and Watson, S. (Eds). *The Palgrave handbook of contemporary heritage research*. Basingstoke: Palgrave Macmillan. pp.414–424.

Schofield, J. (2016). Being autocentric: towards symmetry in heritage management practices. In: Gibson, L. and Pendlebury, J. (Eds). *Valuing historic environments*. Abingdon: Routledge. pp.93–114. [Online]. Available at: doi:10.4324/9781315548449-13 [Accessed 16 November 2020].

Scottish Government. (2023). *Place standard*. [Online]. Available at: https://www.placestandard.scot/ [Accessed 22 August 2021].

Seamon, D. (2020). Place attachment and phenomenology: the dynamic complexity of place. In: Manzo, L. C. and Devine-Wright, P. (Eds). *Place attachment: advances in theory, methods and applications*. 2nd ed. London: Routledge. pp.29–44. [Online]. Available at: doi:10.4324/9780429274442.

Shamai, S. (1991). Sense of place: An empirical measurement. *Geoforum*, 22, pp.347–358. [Online]. Available at: doi:10.1016/0016-7185(91)90017-K.

Sheffield City Council (n.d.). *South Yorkshire historic environment characterisation*. [Online]. Available at: http://www.sytimescapes.org.uk/ [Accessed 14 February 2021].

Smith, L. (2006). *Uses of heritage*. London: Routledge, Routledge. [Online]. [Accessed 21 October 2020].

Social Media Research Group. (2016). *GSR_Social_Media_Research_Guidance_-_Using_social_media_for_social_research.pdf*. [Online]. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/524750/GSR_Social_Media_Research_Guidance_-_Using_social_media_for_social_research.pdf [Accessed 28 August 2021].

Sofaer, J. et al. (2021). Heritage sites, value and wellbeing: learning from the COVID-19 pandemic in England. *International Journal of Heritage Studies*, 27 (11), pp.1117–1132. [Online]. Available at: doi:10.1080/13527258.2021.1955729.

Spennemann, D. H. (2023). *Generative artificial intelligence, human agency and the future of cultural heritage*. SSRN Scholarly Paper, Rochester, NY. [Online]. Available at: doi:10.2139/ssrn.4583327 [Accessed 31 October 2023].

Sporleder, C. (2010). Natural language processing for cultural heritage domains. *Language and Linguistics Compass*, 4 (9), pp.750–768. [Online]. Available at: doi:10.1111/j.1749-818X.2010.00230.x.

Stamp, J. (2013). *Who really invented the smiley face?* [Online]. Available at: https://www.smithsonianmag.com/arts-culture/who-really-invented-the-smiley-face-2058483/ [Accessed 11 February 2022].

Stedman, R. C. (2003). Is it really just a social construction? The contribution of the physical environment to sense of place. *Society & Natural Resources*, 16 (8), pp.671–685. [Online]. Available at: doi:10.1080/08941920309189.

Stefaniak, A., Bilewicz, M. and Lewicka, M. (2017). The merits of teaching local history: Increased place attachment enhances civic engagement and social trust. *Journal of Environmental Psychology*, 51, pp.217–225. [Online]. Available at: doi:10.1016/j.jenvp.2017.03.014.

Stepanchuk, A., Gafurova, S. and Latypova, M. (2020). «Genius loci» as a resource for the development of historical areas of the city. *IOP Conference Series: Materials Science and Engineering*, 890, p.012013. [Online]. Available at: doi:10.1088/1757-899X/890/1/012013.

Stephenson, J. (2008). The cultural values model: an integrated approach to values in landscapes. *Landscape and Urban Planning*, 84, pp.127–139. [Online]. Available at: doi:10.1016/j.landurbplan.2007.07.003.

Swanwick, C. (2004). The assessment of countryside and landscape character in England: An Overview. In: Bishop, K. and Phillips, A. (Eds). *Countryside Planning: New Approaches to Management and Conservation*. Earthscan. pp.253–373.

Taplin, D. H., Scheld, S. and Low, S. M. (2002). Rapid ethnographic assessment in urban parks: a case study of independence national historical park. *Human Organization*, 61 (1), pp.80–93. [Online]. Available at: doi:10.17730/humo.61.1.6ayvl8t0aekf8vmy.

Temple, S. (2019). *Word clouds are lame*. [Online]. Available at: https://towardsdatascience.com/word-clouds-are-lame-263d9cbc49b7 [Accessed 13 February 2022].

Tenzer, M. (2022). Tweets in the Peak: Twitter analysis - the impact of Covid-19 on cultural landscapes. *Internet Archaeology*, (59), Internet Archaeology. [Online]. Available at: doi:10.11141/ia.59.6 [Accessed 19 August 2022].

Tenzer, M. (2023). Social landscape characterisation: a people-centred, place-based approach to inclusive and transparent heritage and landscape management. *International Journal of Heritage Studies*, 30 (3), pp.269–284. [Online]. Available at: doi:10.1080/13527258.2023.2289424.

Tenzer, M. and Schofield, J. (2023). People and places: towards and understanding and categorisation of reasons for place attachment - case studies from the north of England. *Landscape Reserach Journal [Manuscript accepted for publication]*.

Tenzer, M. and Schofield, J. (2024). Using topic modelling to reassess heritage values from a people-centred perspective – applications from the north of England. *Cambridge Archaeological Journal*, (First view). [Online]. Available at: doi:https://doi.org/10.1017/S0959774323000203.

Tenzer, M. et al. (2024). Debating AI in archaeology: applications, implications, and ethical considerations. Internet Archaeology 67. Available at: doi.org/10.11141/ia.67.8.

Thomas, R. M. (2006). Mapping the towns: English Heritage's urban survey and characterisation programme. *Landscapes*, 7:1, pp.68–92. [Online]. Available at: doi:10.1179/lan.2006.7.1.68.

Thwaites, K. (2001). Experiential landscape place: an exploration of space and experience in neighbourhood landscape architecture. *Landscape Research*, 26 (3), pp.245–255. [Online]. Available at: doi:10.1080/01426390120068927.

Thwaites, K. and Simkins, I. (2006). *Experiential landscape: an approach to people, place and space*. 1st ed. London : Routledge. [Online]. Available at: doi:10.4324/9780203462096 [Accessed 17 November 2020].

Tiller, K. (2020). *English local history. An introduction*. 3rd ed. Woodbridge: The Boydell Press.

Tilley, C. (1994). *A phenomenology of landscape. Places, paths and monuments*. Oxford: Berg.

Tilley, C. and Cameron, K. (2017). *An anthropology of landscape: the extraordinary in the ordinary*. London: UCL Press.

Toepoel, V., Vermeeren, B. and Metin, B. (2019). Smileys, stars, hearts, buttons, tiles or grids: influence of response format on substantive response, questionnaire experience and response time. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 142 (1), pp.57–74. [Online]. Available at: doi:10.1177/0759106319834665.

Tompkins, A. (2017). *Oxfordshire historic landscape characterisation project*. Oxfordshire County Council & Historic England. [Online]. Available at: https://www.oxfordshire.gov.uk/residents/environment-and-planning/archaeology/landscape-characterisation [Accessed 23 January 2021].

Tornes, A. (2021). *Enabling the future of academic research with the twitter api*. [Online]. Available at: https://developer.twitter.com/en/blog/product-news/2021/enabling-the-future-of-academic-research-with-the-twitter-api [Accessed 2 November 2023].

Tornes, A. and Trujillo, L. (2022). Enabling the future of academic research with the Twitter API. *Twitter Developer Platform Blog*. [Online]. Available at: https://blog.twitter.com/developer/en_us/topics/tools/2021/enabling-the-future-of-academic-research-with-the-twitter-api [Accessed 11 February 2022].

Torre de la, M. (2002). *Assessing the values of cultural heritage: research report*. Los Angeles: The Getty Conservation Institute., p.125.

tqx94. (2022). *Text-Analytics_LDA*. [Online]. Available at: https://github.com/tqx94/Text-Analytics_LDA [Accessed 14 November 2022].

Traviglia, A. (2022). *Creating our future: creativity and cultural heritage as strategic resources for a diverse and democratic Europe*. In: 2022. Strasbourg . [Online]. Available at: https://rm.coe.int/digital-technologies-including-ai-for-cultural-heritage-in-the-framewo/1680a5ea9b [Accessed 3 November 2022].

Tuan, Y.-F. (1977). *Space and place. The perspective of experience*. Minneapolis: University of Minnesota Press. [Online]. Available at: https://www.academia.edu/19846369/Yi_Fu_Tuan_Space_and_Place [Accessed 2 November 2020].

Tuan, Y.-F. (1979). Thought and landscape. The eye and the mind's eye. In: Meinig, D. W. and Jackson, J. B. (Eds). *The interpretation of ordinary places. Geographic essays*. New York: Oxford University Press. pp.89–102.

Tuan, Y.-F. (1980). Rootedness versus sense of place. *Landscapes*, 24 (1), pp.3–8.

Tuan, Y.-F. (1990). *Topophilia: a study of environmental perceptions, Attitudes, and Values*. New York, United States: Columbia University Press. [Online]. [Accessed 26 December 2020].

Tudor, C. (2012). *An approach to seascape character assessment*. Bristol : Natural England. [Online]. Available at: https://assets.publishing.service.gov.uk/media/5a7e2cb1ed915d74e33f088b/seascape-character-assessment.pdf [Accessed 19 November 2023].

Tudor, C. (2014). *An approach to landscape characteraAssessment*. Natural England. [Online]. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/691184/landscape-character-assessment.pdf [Accessed 27 January 2021].

Tudor, C. (2019). *An approach to landscape sensitivity*. Natural England., p.47.

Turing, A. M. (1950). I.—Computing machinery and intelligence. *Mind*, LIX (236), pp.433–460. [Online]. Available at: doi:10.1093/mind/LIX.236.433.

Turner, S. (2007). Landscape archaeology for the past and future: the place of historic landscape characterisation. *Landscapes*, 8 (2), pp.40–49.

Turner, S. (2018). *Historic landscape characterisation. An archaeological approach to landscape heritage*. [Online]. Available at: doi:10.4324/9781315753423-3 [Accessed 31 January 2021].

Turner, S. and Crow, J. (2009). Silivri and the Thracian hinterland of Istanbul: an historic landscape. *Anatolian Studies*, 59, pp.167–181.

Twitter IR. (2021). *Q4 and fiscal year 2020: Letter to the shareholders*. [Online]. Available at: https://s22.q4cdn.com/826641620/files/doc_financials/2020/q4/FINAL-Q4'20-TWTR-Shareholder-Letter.pdf [Accessed 15 May 2022].

Tyrväinen, L., Mäkinen, K. and Schipperijn, J. (2007). Tools for mapping social values of urban woodlands and other green areas. *Landscape and Urban Planning - Landscape Urban Plan*, 79, pp.5–19. [Online]. Available at: doi:10.1016/j.landurbplan.2006.03.003.

UK Government. (2011). *Localism Act 2011*. [Online]. Queen's Printer of Acts of Parliament. Available at: https://www.legislation.gov.uk/ukpga/2011/20/contents/enacted/data.htm [Accessed 6 February 2021].

UNESCO. (1972). *Convention concerning the protection of the world cultural and natural heritage*. Paris . [Online]. Available at: https://whc.unesco.org/archive/convention-en.pdf [Accessed 19 March 2022].

UNESCO. (1992). *Convention concerning the protection of the world cultural and natural heritage. World Heritage Committee*. Santa-Fe, US . [Online]. Available at: https://whc.unesco.org/archive/1992/whc-92-conf002-12e.pdf [Accessed 19 March 2022].

UNESCO. (1997). *Convention concerning the protection of the world cultural and natural heritage. World Heritage Committee.* Naples, Italy . [Online]. Available at: http://whc.unesco.org/archive/1997/whc-97-conf208-inf4e.pdf [Accessed 19 March 2022].

UNESCO. (2018). *Basic texts of the 2003 convention for the safeguarding of the intangible cultural heritage, 2018 Edition*. [Online]. Available at: https://ich.unesco.org/doc/src/2003_Convention_Basic_Texts-_2018_version-EN.pdf [Accessed 12 November 2020].

UNESCO. (2021a). *Convention concerning the protection of thet world cultural and natural heritage - 1972*. [Online]. Available at: https://whc.unesco.org/en/conventiontext/ [Accessed 12 February 2021].

UNESCO. (2021b). *Impact of the COVID-19 pandemic on world heritage and responses by the Secretariat*. Fuzhou, China .

UNESCO. (2021c). *The operational guidelines for the implementation of the world heritage convention*. Paris . [Online]. Available at: https://whc.unesco.org/en/guidelines/ [Accessed 19 March 2022].

UNESCO. (n.d.). *UNESCO world heritage centre - World Heritage List*. [Online]. Available at: https://whc.unesco.org/en/list/?search=cultural+landscape&themes=4&order=country [Accessed 19 March 2022].

Unicode Consortium. (2020). *The unicodesStandard, Version 13.0*. p.866.

United Nations. (1992). *United Nations sustainable development. Agenda 21*. In: 1992. Rio de Janeiro, Brasil : United Nations. [Online]. Available at: https://sustainabledevelopment.un.org/content/documents/Agenda21.pdf.

University of Edinburgh. (n.d.). *Digimap*. [Online]. Available at: https://digimap.edina.ac.uk/ [Accessed 12 February 2022].

University of Portsmouth. (2017). *Vision of Britain | Daniel Defoe | Letter 8, Part 2: The Peak District*. [Online]. Available at: https://www.visionofbritain.org.uk/travellers/Defoe/30 [Accessed 27 August 2021].

University of Sheffield. (2020). *Roots and futures*. [Online]. Available at: https://www.sheffield.ac.uk/archaeology/research/roots-and-futures [Accessed 14 February 2021].

University of Stirling. (n.d.). *Social value toolkit – Guidance for Heritage Practitioners*. [Online]. Available at: https://socialvalue.stir.ac.uk/ [Accessed 26 January 2022].

Véliz, C. (2021). *Privacy is power: why and how you should take back control of your data: Amazon.co.uk: Véliz, Carissa: 9781787634046: Books*. Corgi.

Verschoof-van der Vaart, W. B. et al. (2020). Combining deep learning and location-based ranking for large-scale archaeological prospection of LiDAR data from The Netherlands. *ISPRS International Journal of Geo-Information*, 9 (5), p.293. [Online]. Available at: doi:10.3390/ijgi9050293.

Vincent, J. (2021). *Facebook bans academics who researched ad transparency and misinformation on Facebook*. [Online]. Available at: https://www.theverge.com/2021/8/4/22609020/facebook-bans-academic-researchers-ad-transparency-misinformation-nyu-ad-observatory-plug-in [Accessed 11 February 2022].

Vries, S. de. (2007). Mapping the attractiveness of the Dutch landscape: a GIS-based landscape appreciation model (Glam2). *Forest Snow and Landscape …*. [Online]. Available at: https://www.academia.edu/3064119/Mapping_the_attractiveness_of_the_Dutch_landscape_a_GIS_based_landscape_appreciation_model_Glam2_ [Accessed 21 September 2022].

de Vries, S. et al. (2013). Measuring the attractiveness of Dutch landscapes: Identifying national hotspots of highly valued places using Google Maps. *Applied Geography*, 45, pp.220–229. [Online]. Available at: doi:10.1016/j.apgeog.2013.09.017.

Wagenaar, P., Rodenberg, J. and Rutgers, M. (2023). The crowding out of social values: on the reasons why social values so consistently lose out to other values in heritage management. *International Journal of Heritage Studies*. [Online]. Available at: https://www.tandfonline.com/doi/abs/10.1080/13527258.2023.2220322 [Accessed 19 June 2023].

Walter, N. (2014). From values to narrative: a new foundation for the conservation of historic buildings. *International Journal of Heritage Studies*, 20 (6), pp.634–650. [Online]. Available at: doi:10.1080/13527258.2013.828649.

Wartmann, F. M., Acheson, E. and Purves, R. S. (2018). Describing and comparing landscapes using tags, texts, and free lists: an interdisciplinary approach. *International Journal of Geographical Information Science*, 32 (8), pp.1572–1592. [Online]. Available at: doi:10.1080/13658816.2018.1445257.

Wartmann, F. M., Koblet, O. and Purves, R. S. (2021). Assessing experienced tranquillity through natural language processing and landscape ecology measures. *Landscape Ecology*, 36 (8), pp.2347–2365. [Online]. Available at: doi:10.1007/s10980-020-01181-8.

Waterton, E., Smith, L. and Campbell, G. (2006). The utility of discourse analysis to heritage studies: the Burra charter and social inclusion. *International Journal of Heritage Studies*, 12 (4), pp.339–355. [Online]. Available at: doi:10.1080/13527250600727000.

Welsh, E. (2002). Dealing with data: using NVivo in the qualitative data analysis process. *Forum: Qualitative Social Research*, 3.

West, S. (2010). Introduction. In: *Understanding heritage in practice*. Manchester : Manchester University Press. pp.1–6.

Williams, D. R. (2000). *Notes on measuring recreational place attachment (unpublished)*. [Online]. Available at: https://www.fs.fed.us/rm/value/docs/pattach_notes.pdf [Accessed 10 April 2021].

Williams, D. R. and Patterson, M. E. (1999). Environmental psychology: mapping landscape meanings for ecosystem management. In: Cordell, H. K. and Bergstrom, J. C. (Eds). *Integrating social sciences and ecosystme menagement: Human dimensioans in assessment, policy and management*. Champaign, IL : Sagamore Press. pp.141–160. [Online]. [Accessed 10 April 2021].

Williams, D. and Vaske, J. (2003). The measurement of place attachment: validity and generalizability of a psychometric approach. *Forest Science*, 49, pp.830–840.

Williamson, T. (2006). Mapping field patterns: a case study from eastern England. *Landscapes*, 7 (1), pp.55–67. [Online]. Available at: doi:10.1179/lan.2006.7.1.55.

Williamson, T. (2007). Historic landscape characterisation: some queries. *Landscapes*, 8:2, pp.64–71. [Online]. Available at: doi:10.1179/lan.2007.8.2.64.

Wilner, A. S. (2018). Cybersecurity and its discontents: artificial intelligence, the internet of things, and digital misinformation. *International Journal*, 73 (2), SAGE, pp.308–316. [Online]. Available at: doi:10.1177/0020702018782496.

de Wit, C. W. (2013). Interviewing for sense of place. *Journal of Cultural Geography*, 30 (1), pp.120–144. [Online]. Available at: doi:10.1080/08873631.2012.745979.

Wood, D. (1992). *The power of maps*, Mappings. New York; London: Guilford Press.

Worthing, D. and Bond, S. (2008). *Managing built heritage: the role of cultural significance*. Oxford : Blackwell. [Online]. Available at: doi:10.1002/9780470697856.

Wuisang, C. E. V. (2014). *Defining genius loci and qualiflying cultural landscape of the Minahasa ethnic community in the North Sulawesi, Indonesia*. Australia : University of Adelaide. [Online]. Available at: https://digital.library.adelaide.edu.au/dspace/bitstream/2440/85188/8/02whole.pdf [Accessed 2 May 2021].

Software and GitHub repositories

Barrett, T., Dowle, M., Srinivasan, A. et al. (2021). *data.table: extension of 'data.frame'*. R package version 1.14.2. Available at: https://cran.r-project.org/web/packages/data.table/index.html.

Bergsma, T. (2021). *csv: read and write CSV files with selected conventions*. R package version 0.6.1. Available at: https://cran.r-project.org/web/packages/csv/index.html.

Bonacchi, C. (2021) *IAR Heritages/R_for_heritage_training_workshop*. Available at: https://github.com/IARHeritages/R_FOR_HERITAGE_TRAINING_WORKSHOP.

Booher, N. (2018) *Qualtrics-google-maps.js*. Available at: https://gist.github.com/njbooher/17ba2845a3490e166c806451afdef387.

De Queiroz, G., Colin, F., Hvitfeldt, E. et al. (2021) *tidytext: text mining using 'dplyr', 'ggplot2', and other Tidy tools*. R package version 0.3.2. Available at: https://cran.r-project.org/web/packages/tidytext/index.html

Eddelbuettel, D. (2021). *digest: create compact hash digests of R objects*. R package version 0.6.29. Available at: https://cran.r-project.org/web/packages/digest/index.html.

Feinerer I, Hornik K (2020). *tm: text mining package*. R package version 0.7-08. Available at: https://CRAN.R-project.org/package=tm.

Fellows I (2018). *wordcloud: word clouds*. R package version 2.6. Available at: https://CRAN.R-project.org/package=wordcloud.

Hornig, K. (2018) *NLP: natural language processing infrastructure.* R package version 0.2-0. Available at: https://cran.r-project.org/web/packages/NLP/index.html

Hutto, C.J. & Gilbert, E.E. (2020). *VADER: a parsimonious rule-based model for sentiment analysis of social media text*. Version 3.2.2. Available at: https://pypi.org/project/vaderSentiment.

Jones, T., Doane, W. and Attbom, M. (2021). *textmineR: functions for text mining and topic modeling.* R package version 3.0.5. Available at: https://CRAN.R-project.org/package=textmineR.

Jones, T. W. (2021). *Topic modeling*. Available at: https://cran.r-project.org/web/packages/textmineR/vignettes/c_topic_modeling.html.

Bouchet-Valat, M. (2020). *SnowballC: Snowball stemmers based on the C 'libstemmer' UTF-8 library*. R package version 0.7.0. Available at: https://cran.r-project.org/web/packages/SnowballC/index.html.

Müller, K. (2021). *Tidiverse/tibble*. R package version to 3.2.5. Available at: https://cloud.r-project.org/web/packages/tibble/index.html.

Neuwirth, E. (2022). *RColorBrewer: ColorBrewer Palettes*. R package version 1.3-1. Available at: https://cran.r-project.org/web/packages/RcolorBrewer/index.html

Omkar, P. (2019). *Emosent-py*. Version 0.1.6. Available at: https://pypi.org/project/emosent-py

Ooms, J. and McNamara, J. (2021). *writexl: export data frames to Excel 'xlsx' Format*. R package version 1.4.0. Available at: https://cran.r-project.org/web/packages/writexl/index.html

Python Software Foundation (2019). *Python language reference.* Version 3.8.10. Available at http://www.python.org.

R Core Team (2021). *R: A language and environment for statistical computing*. Version 4.1.1. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org.

Spannbauer, A. (2019). *lexRankr: extractive summarization of text with the LexRank algorithm*. R package version 0.5.2. Available at: https://cran.r-project.org/web/packages/lexRankr/index.html.

Tyler, R. (2018). *textstem: tools for stemming and lemmatizing text*. R package version 0.1.4. Available at: https://cran.r-project.org/web/packages/textstem/index.html.

tqx94. (2022). *Text-analytics LDA*. Available at: https://github.com/tqx94/Text-Analytics_LDA.

Ushey K, Allaire J, Tang Y (2022). *Reticulate: interface to 'Python'.* R package version 1.26. Available at: https://CRAN.R-project.org/package=reticulate.

Wickham H, Chang, W., Henry, L., et al. (2021). *ggplot2: Create elegant data visualisations using the grammar of graphics*. R package version 3.3.5. Available at: https://cran.r-project.org/web/packages/ggplot2/index.html.

Wickham H, François R, Henry L, Müller K et al. (2022). *dplyr: a grammar of data manipulation*. R package version 1.0.10. Available at: https://cran.r-project.org/web/packages/dplyr/index.html.

Wickham, H. (2022). *stringr: simple, consistent wrappers for common string operations*. R package version 1.4.1. Available at: https://CRAN.R-project.org/package=stringr.

Wickham, H., (2021) *Tidyverse: easily install and load the 'Tidyverse'*. R package version 1.3.1. Available at: https://cran.r-project.org/web/packages/tidyverse/index.html.

Wickham H (2018). *reshape: flexibly reshape data*. R package version 0.8.8. Available at: https://cran.r-project.org/web/packages/reshape/index.html.

Wickham H, Henry L (2021). *purrr: functional programming tools*. R package version 0.3.5, Available at: https://cran.r-project.org/web/packages/purrr/index.html.

Wickham H, Vaughan D, and Girlich M (2022). *tidyr: tidy messy data.* R package version 1.2.0, Available at: https://CRAN.R-project.org/package=tidyr

Wickham H, Bryan J (2022). *readxl: read Excel files*. R package version 1.4.1. Available at https://CRAN.R-project.org/package=readxl.

QGIS Development Team (2021). *QGIS Geographic Information System*. Open Source Geospatial Foundation Project. QGIS version 3.22LTR (Białowieża). Available at: http://qgis.osgeo.org.

A. Appendices for Chapter 3:

Appendix 1

The following code block shows one search query parameter set for the second study period.

```
query1 <- build_query(
  query = c('#peakdistrict', '#PeakDistrictNationalPark',
            '#peakdistrictwalks', 'Peak District',
            'Peak District National Park'),
  is_retweet = FALSE,
  remove_promoted = TRUE,
  lang = 'en',
  exclude = c('recruitment', 'recruiting', '@WhaleyChronicle',
  '@weatherwhaley', '#WeAreWorkingForYou', '@etsyUK', '@etsy',
  'price', '#ukgetaway', '@etsy')
)

get_all_tweets(
    query1,
    start_tweets = '2020-05-23T00:00:59Z',
    end_tweets = '2020-05-25T23:59:59Z',
    data_path = 'tweetdata-20200525',
    bind_tweets = FALSE,
    n = 2000
  )
```

Appendix 2

The renv dependency file can be found in the online version of this paper and archived at: https://intarch.ac.uk/journal/issue59/6/renv.lock

Appendix 3

Find Places - Manual Routine

Finding Places

In this part of the process, we are looking to find places based on a curated place list that was created from OS data, Historic England datasets and Digimap (Edina). This workbook will identify places in a data frame of cleaned texts and create a location frequency table for visualisation in GIS.

Read data frame of texts

```
library(readxl)

tweetdf <- read_xlsx(paste("tweetsdfcleaned_", startdate, ".xlsx", sep = "
"))
```

Read the place list (gazetteer)

```
placeSearchList <- read.csv("../List_POI_places.csv")$Places
```

Place matching routine

This is a simple routine that looks for matches of single words and matches a multi-word location when part of it is matched and the other part is not in the place-match stop-word list (which is hand curated so (a) work in progress, and (b) a potential source of bias/error). For example, “Mam” will not match “Mam Tor”, but “Sir William” will match “Sir William Hill”. This has a risk of both false positives and positives for sloppy texts, but that seems to be the best way to add some fuzzyness to the matching.

Outlook: use Machine Learning (ML) or Natural Entity Recognition (NER), currently these do not work for locations in the UK, since the training data is US centric. Particularly, this routine will support for NER for specific places, e.g. national parks.

Clean the tweet text

We need to clean the social media further. We use `tm` and a corpus to do that since it is easier that way (and because doing it on a dataframe caused nested weirdness).

```
library(tm)

corpus <- Corpus(VectorSource(tweetdf$cleantext))

removeHashTags <- function(x) gsub('#', ' ', x)
removeReferences <- function(x) gsub('@', ' ', x)
removeAMP <- function(x) gsub('&amp;', ' ', x)
```

```
removeEllipses <- function(x) gsub('\U{2026}', ' ', x)
removeEllipses2 <- function(x) gsub('[\\.]+', ' ', x)
removePunc <- function(x) removePunctuation(x, ucp = TRUE)
removeQuotes <- function(x) gsub('’', ' ', x)
# removeURL <- function(x) gsub('(f|ht)(tp)(s?)(://)(.+)[.|/]([a-zA-Z0-9_-
]+)\\b', ' ', x)

# corpus <- tm_map(corpus, removeURL)
corpus <- tm_map(corpus, removeHashTags)
corpus <- tm_map(corpus, removeReferences)

corpus <- tm_map(corpus, removeAMP)
corpus <- tm_map(corpus, removeEllipses)
corpus <- tm_map(corpus, removeEllipses2)
corpus <- tm_map(corpus, removePunc)
corpus <- tm_map(corpus, removeQuotes)

corpus <- tm_map(corpus, stripWhitespace)

tweetdf$cleanhashtext<- corpus$content
```

Define the place search function

```
matchPlaces <- function(text) {
  #foundPlaces = vector()
  foundPlaces <- ""
  for (place in placeSearchList) {
    # placeWord matches with word boundary markers "\b" - double escapes!
    placeWord <- paste("\\b", place, "\\b", sep="")
    wordMatch <- any(grep(placeWord, text, ignore.case = TRUE))
    # placeMerged matches concatenated place name with word boundary marke
rs
    placeMerged <- paste("\\b", gsub(" ", "", place), "\\b", sep="")
    mergeMatch <- any(grep(placeMerged, text, ignore.case = TRUE))

    # both wordMatch and mergeMatch are just booleans now, to make the
    # if statement neater:
    if (wordMatch | mergeMatch) {
      #print(place)
      if (nchar(foundPlaces) == 0) {
        foundPlaces <- place
      }
      else {
        foundPlaces <- paste(foundPlaces, place, sep=", ")
      }
    }
  }
  #if (length(foundPlaces) > 0) {
    return(foundPlaces)
  #}
  #else {
  #  return("")
```

```
  #}
}
```

Note: I liked the idea of a vector, but that caused nested data frame columns (tibble in data frame, when using lapply - actually lapply/mapply that does that, so I still need to understand how to work with them properly).

Test:

```
matchPlaces("This is a text about Stanage Edge, the A625, and Mam Tor and
with a hash tag about #LadyBowerReservoir, written @winhill")
```

Seems to work.

For loop method

We try a traditional *for* loop and assign that result to the column `placesFound`:

```
tweetdf[ , 'placesFound'] <- NA

system.time(
for (i in 1:nrow(tweetdf)) {
  tweetdf[i, 'placesFound'] <- matchPlaces(tweetdf[i, "cleanhashtext"])
}
)

tweetdf[, 'placesFound']
```

Have a closer look at the data

It is easier to see what is going on in a simpler dataframe

```
placesdf <- tweetdf[,c('text', 'cleanhashtext', 'placesFound')]

library(writexl)
write_xlsx(placesdf, "placesdf.xlsx")
write_xlsx(tweetdf, paste("tweetsdflocations_", startdate, ".xlsx", sep =
""))

saveRDS(tweetdf, file = paste("tweetsdflocations_", startdate, ".RDS", sep
= ""))
```

Location frequency, automated process

I changed this to read the RDS file, since reading the Excel file changes some characters, e.g.the apostrophe `'` comes out of Excel as `’`, which is not the same and does not match when trying to match the coordinates later!

```
notationdf <- readRDS(paste("tweetsdflocations_", startdate, ".RDS", sep =
""))
library(tm)
```

```
corptextloc <- notationdf$placesFound
corptextloc <- Corpus(VectorSource(corptextloc))
removeblanks <- function(x) gsub(' ', '_', x)
corptextloc <- tm_map(corptextloc, removeblanks)
replaceblank <- function(x) gsub(',_', ' ', x)
corptextloc <- tm_map(corptextloc, replaceblank)

dtmlocs  <- TermDocumentMatrix(corptextloc)
inspect(dtmlocs[1:10,1:10])

library(data.table)
mydtmlocs <- as.matrix(dtmlocs)
a <- sort(rowSums(mydtmlocs), decreasing=TRUE)
Freqlocs <- data.frame(place = gsub("_", " ", names(a)), count =a)
head(Freqlocs,10)  #this sorts the first 10 word for the visualisation on
freq
```

Add Location Data from the Places List

First create the x and y columns in `Freqlocs`:

```
Freqlocs$x <- NA
Freqlocs$y <- NA
```

Read the place search dataframe:

```
placeSearchdf <- read.csv("../List_POI_places.csv")
```

We can use the built in query function to select the right rows from the place search dataframe:

```
placeSearchdf[placeSearchdf$Places == "Mam Tor", ]
```

A test to see if that works with the `tolower` function as well:

```
placeSearchdf[tolower(placeSearchdf$Places) == "mam tor", ]
```

It works! So we can then get the coordinates like this (`jacob's ladder` is the one that wasn't recognised when reading in the `xlsx` file, but it does work when using the `RDS`):

```
placeSearchdf[tolower(placeSearchdf$Places) == "jacob's ladder", 'X']
```

We can now add the x and y values to the places found.

```
Freqlocs$x <- mapply(function(loc) placeSearchdf[tolower(placeSearchdf$Pla
ces) == loc, 'X'], Freqlocs$place)
Freqlocs$y <- mapply(function(loc) placeSearchdf[tolower(placeSearchdf$Pla
ces) == loc, 'Y'], Freqlocs$place)
```

And save the location dataframes with the coordinates:

```
saveRDS(Freqlocs, "Freqlocs.RDS")
write_xlsx(Freqlocs, "Freqlocs.xlsx")
write.csv(Freqlocs, file="Freqlocs.csv")
```

These can now be read into GIS to visualise hotspots of places that were mentioned in the social media or survey texts.

Additional information

As the API does not exist any longer, after changes to the X's access and research policies, the code for the analysis is not archived and made publicly available.

B. Appendices for Chapter 4:

Supplementary material 1

International documents on the conservation and management of heritage with key statements to social values and participation.

Document	Alias	Year	Key statements
Amsterdam Declaration adopted at the Congress on the European Architectural Heritage ICOMOS (1975)	Amsterdam Declaration	1975	"Local authorities should improve their techniques of consultation for ascertaining the opinions of interested parties on conservation plan and should take these opinions into account from the earliest stages of planning. As part of their efforts to inform the public the decisions of local authorities should be taken in public using a clearly understandable language so that the local inhabitants may learn, discuss and assess the ground for them. Meeting places should be provided in order to enable members of the public to consult together...complementary proposals or alternatives put forward by groups or individuals should be considered as an important contribution to planning."
The Australia ICOMOS Guidelines for the Conservation of Places of Cultural Significance (ICOMOS 1979)	Burra Charter	1979	No reference to public participation or people/place connection, revisions 1981, 1988, 1999, 2013 (current version)
The Australia ICOMOS Charter for Places of Cultural Significance ICOMOS (1999)	Burra Charter	1999	Groups and individuals with associations with a place as well as those involved in its management should be provided with opportunities to contribute to and participate in understanding the cultural significance of the place. Where appropriate they should also have opportunities to participate in its conservation and management.
Council of Europe: European Landscape Convention (Council of Europe 2000)	Florence Convention	2000	"One of the major innovations of the European Landscape Convention is the definition of "landscape quality objectives", meaning, for a specific landscape, the formulation by the competent authorities of the aspirations of the public with regard to the landscape features of their surroundings. No longer the preserve of experts, landscape is now a policy area in its own right." (https://www.coe.int/en/web/landscape/the-european-landscape-convention)
Convention on the Value of Cultural Heritage for Society (Council of Europe 2005)	Faro Convention	2005	"Emphasizes the important aspects of heritage as they relate to human rights and democracy. It promotes a wider understanding of heritage and its relationship to communities and society. The Convention encourages us to recognize that objects and places are not, in themselves, what is important about cultural heritage. They are important because of the meanings and uses that people attach to them and the values they represent." (https://www.coe.int/en/web/culture-and-heritage/faro-convention)
Conservation Principles, Policies and Guidance Historic England (2008)	Conservation Principles	2008	''everyone should be able to participate in sustaining the historic environment, by having the opportunity to contribute his or her knowledge of the value of places, and to participate in decisions about their future, by means that are accessible, inclusive and informed"

The Australia ICOMOS Guidelines for the Conservation of Places of Cultural Significance ICOMOS (2013)	Burra Charter	2013	"Places may have a range of values for different individuals or groups. Conservation, interpretation and management of a place should provide for the participation of people for whom the place has significant associations and meanings, or who have social, spiritual or other cultural responsibilities for the place."

Supplementary material 2

Examples of stories provided by survey participants, which give an insight of the type of reason for the connection between people and places.

Category	Subcategory	Place	Story
Green Space		Deep Dale, Monyash, PDNP	"Visit regularly- nature, walks, countryside, camping all in this rich area where still meadows and sympathetic farming methods and wide range of wild plants and life, nature reserves".
		Longshaw Estate, PDNP	"An area of natural beauty owned and managed by the national trust. I volunteer here with others working under the direction of rangers on the estate. the work is sympathetic management of local environment to promote and sustain the natural ecology and reduce/repair wear and tear, erosion and damage from people using and enjoying the area."
	Health	Loxley Valley, Sheffield	"Now that I have retired, I am appreciating it more and more; during lockdown it was the best place to be! We could just walk out of the house and follow the many footpaths up the valley to the moors above Ladybower."
		Botanical Gardens, Sheffield	"Again, this is very close to where we live, and we almost consider it as our garden. It was also very important for meeting our family during the pandemic."
		Endcliffe Park Cafe, Shefifeld	"Cafe in a small pavilion that forms a central hub in Endcliffe Park - lots of time spent there with friends and family pre-COVID and a good place to meet outside for a socially distanced takeaway cuppa during COVID"

		Local parks off Jaunty Lane to Birley, Sheffield	“Covid Lockdown Sanctuary in 2020, didn't know some of it existed. Still walking in the area now.”
Communal	Spiritual	Hathersage Road layby, PDNP	“My husband used to drive out and park in the layby during his lunch hour to read. His ashes are buried on the moor around here. We come out and park up to think about him.”
		Stanage Edge and Scots Pine Tree, PDNP	“Some of my dad’s ashes are scattered on the top of the edge - a more dramatic thing than we'd expected as the rain and wind were far worse on the top than in the valley - some ashes blew back on me! The rest I scattered round the foot of a pine just off the path that leads to the carpark, as we came back down. Here by contrast it was quiet and completely still. A profound and moving time of completely contrast weather that he as a geographer would have been equally fascinated by.”
		Above Bretton Brook, PDNP	“Scene of many walks, picnics etc Used to be a magic place till it got a bit overgrown. Would like ashes scattered there.”
		Bailey Hill, High Bradfield, PDNP	“A scheduled monument behind the church. Much frequented by villagers and visitors. My neighbour & good friends ashes are scattered here along with his wife.”
	Arts & Culture	Crucible Theatre, Sheffield	“What a fabulous place. And the Crucible. We’ve seen so many amazing plays here. A real Jewell.”
		Sheffield City centre	“The museums in Sheffield including Kelham, Abbeydale, Weston Park and the museums and galleries in the city centre keeping our heritage alive.”

Supplementary material 3

Questionnaire questions:

General Questions

1. I am a resident of… (options 1. Sheffield, 2. Peak District National Park)
2. Where did you find this survey?
3. What is your age?
4. Nationality?
5. Ethnicity?
6. Identity?
7. Gender?
8. Highest level of education?
9. Please enter your postcode.
10. How many years have you lived here (in total)?
11. Reasons for living here (Sheffield/Peak District)?
12. How likely is it that you relocate in future?
13. Do you work in archaeology, heritage, conservation, or any related field of the historic environment?
14. Are you a member of a volunteer organisation?

Special places

1. Your special places
   Please tell me about the places, landmarks, buildings, landscapes or objects in Sheffield and the Peak District that have personal meaning to you, a place that you feel a special attachment to. This can be a place in nature, a forest, a tree, a river, rolling fields, or wider landscapes. It can also be a building, a street, green spaces, parks, a canal, a favourite walk or recreational areas, an old streetlight or a traditional phone box, areas with graffiti or street art. If there is a special place or object that makes you feel angry, anxious, or uneasy, this counts as a special place as well.

   How much does the history and the character of the place play a role in your sense of belonging and identity?

   Please enter up to 5 places in both Sheffield and the Peak District, irrespective of where you live.

2. Location identified either on a map or textual entry.
3. How much change would you accept in this place?
4. Do you have photos, drawings, sketches, or other images of the place you want to share? (This was repeated for four more places)

The Historic Environment

1. How would you define local heritage in your area?
2. In three words/expressions: what represents Sheffield's unique character?
3. In three words/expressions: what represents the Peak District's unique character?
4. Please rate the following statements (Likert scale):
   - The history and the past are…
   - The local heritage is …
   - Nature and recreation are …
   - People and community is …
   - Cultural diversity is …
   - The material and fabric of places are …
   - The essence/spirit of a place is …
   - Traditions and preservation are …
   - Change and development are …
5. Are there any special traditions, traditional skills, historic events, habits that are important to you?
6. Why should we preserve our local heritage?
7. What can we do to support the local area and heritage?
8. Would you be interested in a follow-up interview?
9. Email address.

C.Appendices for Chapter 5:

Supplementary material 1

The list of interview partners, their place of residence in case it is located in the PDNP, and the occupation is detailed below. Note: the interview partners waived their right to anonymity as the cases were particular to the situation and location. Interviewees included:

- the longest serving drystone waller in the Peak District (Trevor, Interview 1)
- farmer and director of the Dove Valley Centre (Elspeth, Interview 2)
- farmer's daughter and initiator of the Waterhouse Farm Barn Restoration project (Julia, Interview 3)
- Lockerbrook Outdoor Centre Manager (Jo, Interview 4)
- Community volunteer at the “plague village” Eyam (Joan, Interview 5)
- farmer and Hardhurst Farm Camping manager (Sue, Interview 6)
- retired farmer, artist, and holiday cottage manager (Sue, Interview 7)
- Breedon Group, Hope Cement Plant Sustainability Officer (Spencer, Interview 8)
- Hope Valley Climate Action group coordinator (David, Interview 9)
- Head of “Moors for the future” partnership (Chris, Interview 10)

Supplementary material 2

Topics and keywords resulting from Topic Modelling. Labels are created based on keyword bi-grams.

Topic	Label 1	Label 2	Top Terms	Manual Topic
t_1	stone_wall	dry_stone	stone, wall, teach, spin, good, stone_wall, spin_dry, dry, give, dry_stone, job, time, derbyshire, build, fit, enjoy, trevor, pass, limestone, learn, nice, preserve, people, carl, thing, give_time, art, lose, shape, win, tradition, quarry, start, bite, earn, gentleman, technique, lady, hundred, product, middle, realise, change, advise, bridge, build_gap, derbyshire_branch, enjoy_job, gritstone, jigsaw	17
t_2	cement_work	twenty_year	cement, quarry, people, plant, carbon, park, good, cement_work, energy, make, term, work, uk, reduce, business, cement_plant, quarry_quarry, sustainability, breedon, company, product, sustainable, carbon_emission, emission, transport, key, source, element, safety, concrete, health, side, group, build, conscious, ireland, reaction, uk_ireland, work_cement, alternative, people_hate, tonne, engineer, fuel, stakeholder, hate, ruin, light, environment, run	11
t_3	young_people	local_authority	group, young, young_people, people, work, activity, talk, centre, project, lot, louisa, set, school, start, part, education, youth, charity, content, outdoor, programme, team, write, involve, end, support, run, voice, pandemic, benefit, lead, explore, meet, amaze, create, john, aim, child, ambassador, connect_young, group_group, john_muir, maker, muir, offer, platform, award, join,	19

			development, engage	
t_4	national_park	good_place	place, walk, good, bite, north, interest, dale, big, favourite, yorkshire, boundary, edge, hill, talk, love, park, time, guess, area, favourite_place, kid, wild, moor, run, swim, northern, west_yorkshire, good_place, west, cycle, study, generally, weekend, special, great, bradford, occasionally, yorkshire_dale, hour, stanage, buxton, river, visit, lovely, top, nice, find, compromise, good_compromise, northern_edge	23
t_5	good_good	national_park	farm, good, move, small, buy, side, brother, love, husband, family, caravan, mother, good_good, feel, life, time, farm_side, farm_farm, hate, generation, nice, live, campsite, small_farm, camp, father, discover, sell, absolutely, realise, difficult, bring, cost, farm_good, farmer_good, good_farm, hate_hate, husband_family, lamb, nowadays, family_farm, mortgage, friendly, funny, uncle, employ, god, sense, remember, hill	10
t_6	story_place	young_people	history, people, place, story, important, learn, give, build, time, interest, bite, happen, question, site, story_place, hear, sit, past, history_place, century, bit, part, long, sort, crash, give_bite, place_history, read, case, form, point, give_story, place_sort, people_give, tudor, add, rock, church, understand, lot, basinstone, crash_site, history_happen, hollow, industrial_history, lot_bit, packhorse,	20

			packhorse_route, place_important, place_story	
t_7	national_park	national_trust	park, national, national_park, people, live, visitor, bite, good, heritage, important, sort, context, front, create, find, park_authority, people_park, pressure, half, authority, huge, pay, people_live, move, life, effect_people, park_national, park_visitor, apply, awful, car_park, enormous, partnership, massive, effect, negative, dovedale, car, protect, business, suppose, landscape, work, attract, attract_people, farmyard, find_national, park_act, park_move, pay_park	7
t_8	drystone_wall	stone_wall	wall, build, put, stone, drystone, drystone_wall, style, picture, job, important, build_wall, trevor, feature, show, money, good, bee, back, tradition, day, waller, hide, find, people, lichen, put_wall, superstitious, wall_wall, rabbit, type, garden, wildlife, nice, run, olden, olden_day, put_stone, shoe, wall_build, smoots, list, mark, stand, rare, high, preserve, realise, site, leave, work	13
t_9	place_place	sort_place	place, feel, make, live, child, home, call, back, location, quiet, garden, relax, door, move, daughter, peace, flat, place_place, son, ah, stop, time, drop, make_feel, place_home, place_location, ground, miss, people_feel, pottage, law, safe, boy, enjoy, city, environment, nice, leave, talk, life, run, part, landscape, thing, farmhouse, feel_home, feel_safe,	27

			group_call, home_home, home_landscape	
t_10	national_trust	make_thing	thing, make, interest, understand, good, find, member, make_thing, society, thing_good, train, change, time, national_trust, influence, human, year, affect, chatsworth, experience, trust, realise, stuff, sort, observe, deeply, frustrate, organisation, interest_thing, decision, sustainable, light, learn, affect_life, benevolent, contrast, good_make, member_national, planner, find_interest, tenant, invest, fight, mistake, panel, ghost, encourage, read, god, moment	6
t_11	climate_change	blanket_bog	water, climate, change, climate_change, bog, blanket, blanket_bog, degrade, peat, condition, upland, moor, future, important, quality, moor_future, emergency, world, south, climate_emergency, gather, authority, tree, back, active, record, provide, head, restore, sustainable, bog_condition, drink_water, local_authority, dark_south, ecosystem, drink, grind, bad, idea, term, run, work, thing, active_blanket, archaeological_record, degrade_upland, draught, ecosystem_service, water_gather, achieve	2

t_12	lime_stone	wildlife_trust	thing, kind, side, stuff, site, park, guy, stone, local, lime_stone, basically, trust, lime, natural, plan, back, side_thing, stakeholder, wildlife_trust, conversation, give, kind_stuff, mechanical, derbyshire_wildlife, employee, complex, thing_thing, easy, age, volunteer, pretty, support, pole, simple, employ, replace, ruin, rock, management, early, skill, wildlife, biodiversity, form, week, derbyshire, nature, project, run, year	17
t_13	point_view	plant_tree	good, point, people, biodiversity, plant, area, view, environment, tree, point_view, species, difficult, start, landscape, sort, contribution, make, resource, put, bird, nature, problem, plant_tree, population, middle, wildlife, give, important, difficult_people, tree_good, survive, evidence, woodland, rare, country, wilderness, hill, arena, people_view, plant_good, sphagnum, black, bleaklow, finance, ecosystem, depend, reason, habitat, state, full	3
t_14	sheep_farm	tree_grow	farm, farmer, land, big, sheep, change, grow, government, tree, gate, rewilding, pay, bite, sheep_farm, lot, happen, tree_grow, bad, idea, form, interest, work, gate_place, subsidy, rent, common, generally, pretty, economically, machinery, marginal, small_gate, subsidise, differently, income, post, develop, turn, character, hedge, case, job, term, small, feel, ability, acidification, big_big, change_character, common_vision	12

t_15	lime_kiln	people_walk	wood, lime, field, tree, valley, place, walk, top, side, landscape, amaze, beautiful, ilam, green, light, kiln, river, lime_kiln, extraordinary, hill, dovedale, manifold, gorge, hole, woodland, sell, sheep, land, big, wall, people_walk, stun, astonish, manifold_valley, bottom, leave, bite, arch, grand, ilam_park, secrete, secrete_place, side_valley, uh, bluebell, opposite, beechenhill, stone, part, back	26
t_16	hay_meadow	drystone_wall	meadow, hay, plant, field, sheep, hay_meadow, diverse, grass, graze, hedge, flower, rich, habitat, create, grassland, winter, manage, grow, wildflower, insect, control, disappear, full, boundary, project, start, grassland_project, plant_hedge, range, rich_diverse, seed, unspoiled, wildflower_meadow, cattle, cow, replace, diversity, bird, rare, hard, road, call, land, part, ash, ash_tree, create_hay, creature, field_boundary, flora	1
t_17	pilsbury_castle	put_back	house, back, find, live, buy, valley, fireplace, paul, set, land, decide, hall, piece, year, medieval, wale, fall, child, family, cave, belong, win, fact, put, history, cellar, pottery, put_back, family_live, ruin, hold, high, putting, explore, teach, feel, auction, cellar_house, dispute, horrobyn, lintel, medieval_house, teacher, include, tudor, discover, develop, amaze, lovely, leave	20
t_18	farm_wilderness	find_lovely	find, story, thing, dog, mine, tip, people, lovely, farm, website, bowl, book, garden, dovedale, wilderness, ochre, gentle, room, craft, shop, character, stay, walk, find_lovely, pot, wash, medium, minute, movement, pull, lady, pandemic, word, love, interest, allegedly, alport, alport_castle, bowl_tip, castle_farm, dog_bowl, donkey, ochre_mine,	24

			pull_people, room_tip, sheep_wash, suffragette, tip_bowl, tip_dog, trade	
t_19	good_dress	good_good	good, year, day, longnor, skill, make, tradition, week, race, horse, repair, covid, start, sort, time, wake, beautiful, happen, carnival, dress, good_dress, pub, sunday, basket, event, good_good, locally, important, happen_year, september, meet, term, bite, thing, originally, place_land, saint, wake_week, year_covid, celebrate, weave, twelve, craft, month, hedge, rural, centre, early, dark, involve	13
t_20	twenty_year	year_ago	year, people, twenty, leave, twenty_year, big, part, impact, year_ago, ago, good, point, material, thirty, world, end, bring, run, driver, late, move, time, raw, fourty, plague, wrong, breedon, hundred, close, change, sort, people_leave, raw_material, till, twenty_thirty, big_impact, fourty_year, independent, factory, knock, traffic, company, century, good_good, kid, hard, job, manage, hope, start	11
t_21	national_trust	young_people	people, work, good, happen, thing, change, make, fund, future, money, people_good, people_work, public, hard, place, numb, local_people, protect, beautiful, big, basket, good_thing, deal, people_people, team, conversation, trust, heritage, stuff, local, history, interest, future_good, role, change_happen, good_big, continue, thing_happen, hold, national_trust, grant, hedge, case, experience, stop, landscape, fund_public, make_happen,	15

			role_make, view_people	
t_22	mum_dad	dad_live	dad, thing, live, bear, mum, road, call, work, mum_dad, marry, dad_live, thousand, pass, close, mam, mam_tor, tor, castleton, kid, train, village, good, rail, coal, bottom, buxton, open, quarry, winhill, hathersage, track, edale, route, mind, hope, side, valley, year, bear_mind, inherit, phillip, pike, topley, topley_pike, appeal, lorry, ridge, museum, seventy, line	21
t_23	grow_oat	king_consort	animal, human, top, shelter, day, castleton, play, good, garland, king, water, village, provide, winter, call, band, museum, break, spring, anymore, hill, big, time, consort, king_consort, shoulder, handle, ride, weather, girl, cover, mile, horse, thousand, shape, fall, barn, nice, long, put, important, animal_sheep, band_king, bell, castleton_garland, festival, flower_king, garland_castleton, good_animal, human_animal	9
t_24	lump_bump	incredibly_important	paint, landscape, thing, story, people, colour, land, make, show, church, bump, lump, lump_bump, clump, corner, incredibly, line, completely, absolutely, person, future, field, important, walk, incredibly_important, story_story, draw, art, environment, happen, clump_tree, landscape_landscape, landscape_shape, trough, landscape_environment, people_paint, trail, panel, inhabit, worry, shape, god, picture, tree,	20

			community, church_build, church_people, community_paint, farm_trail, folk_art	
t_25	people_live	family_live	live, family, sort, back, grow, connection, suppose, feel, place, grandfather, root, bite, generation, people_live, great, child, farm, neighbour, long, life, part, start, childhood, research, play, find, family_live, drive, realise, grandmother, memory, live_live, estate, week, lot, kind, call, tree, family_member, family_tree, grandfather_live, great_grandfather, maternal, household, live_family, sheen, document, mill, mom, atmosphere	21
t_26	national_park	people_visit	heritage, thing, visit, local, sort, remember, landscape, stuff, advantage, personal, food, talk, local_heritage, tourism, england, walk, back, good, lockdown, father, guess, feature, question, run, people_visit, idea, natural, sense, environment, happen, bite, work, figure, suddenly, man_make, mention, straight, foot, real, aspect, strong, front, man, mind, farmer, life, capture_aspic, disconnect, father_father, food_local	4

t_27	stewardship_scheme	tree_plant	scheme, grouse, shoot, good, small, area, people, term, tree, mindset, moor, plant, number, grant, support, run, stewardship, stewardship_scheme, tree_plant, access, environmentally, grouse_moor, stock, wildlife, money, sensitive, sensitive_area, basic, payment, produce, fence, environmental, plant_tree, population, food, rewilding, manage, farm, basic_payment, environmentally_sensitive, gamekeeper, number_grouse, people_access, predator, small_number, stock_rate, rate, approach, wet, feed	8
t_28	wildlife_trust	sort_thing	people, thing, sort, community, group, day, church, village, money, base, die, sort_thing, lot, local, hand, field, apple, ghost, people_thing, set, call, run, thing_people, event, repair, pay, form, base_place, make_money, press, maintain, wildlife_trust, rest, paul, watch, play, river, trust, wildlife, derbyshire, amaze, school, put, young, place, work, community_support, dip, event_group, family_lie	18
t_29	year_ago	ten_year	year, ago, year_ago, castle, find, ten, ten_year, project, child, pilsbury, person, pilsbury_castle, goodness, valley, dig, excite, fund, end, fact, happen, family, big, norman, million_year, extraordinary, protect, open, local, people, lottery, lottery_fund, de, sea, hartington, finish, spring, book, million, absolutely, involve, great, school, water, day, archaeologist, attraction, big_project, castle_norman, de_ferrers, family_build	22

t_30	hope_valley	work_farm	work, good, thing, hope, valley, support, hope_valley, part, bite, bring, club, community, business, picture, life, year, leave, golf, support_local, tie, circle, local, interest, area, people, football, golf_club, blue, blue_circle, people_bite, fight, work_work, happy, estate, challenge, thing_good, involve, stuff, plant, important, identity, people_hope, sense_community, thing_alive, thing_plant, work_life, bet, gym, historically, popular	15
t_31	bamford_edge	monsal_head	love, beautiful, sort, time, lovely, run, valley, walk, dark, edge, word, nice, beauty, monsal, white, wild, call, busy, route, head, bamford, special, space, view, road, kind, bamford_edge, bleak, gorgeous, monsal_head, cold, reservoir, fact, suppose, water, place, good_beautiful, lovely_walk, love_love, ancient, wind, stanage, diverse, person, south, beautiful_busy, beautiful_love, beautiful_wild, diverse_diverse, gorgeous_love	25
t_32	plan_permission	farm_building	good, building, build, plan, barn, work, suppose, make, stop, permission, plan_permission, time, awareness, level, restore, area, roof, cottage, form, place, farm_building, issue, perspective, feature, animal, important, bite, good_build, unusual, inside, burn, put, thing, building_good, demonstration, redundant, redundant_farm, typical, tunnel, countryfile, cow, room, street, anymore, modern, negative, countryside, ten, show, money	14

t_33	people_live	holiday_home	people, house, live, school, home, move, time, day, holiday, stay, life, year, young, remember, friend, people_live, family, community, property, child, village, make, work, good, watch, countryside, man, locally, buy, big, holiday_home, tideswell, big_house, granddad, young_people, local, part, cottage, bite, thing, family_home, good_friend, granddad_live, grandma, young_man, bank, empty, good_time, privilege, traffic	7
t_34	young_people	woodcraft_folk	people, place, connect, space, city, nature, save, kind, woodcraft, find, sense, back, woodcraft_folk, culture, folk, positive, diversity, speak, perspective, volunteer, money, lockerbrook, save_space, connection, bring, guess, rural, understand, love, young, land, derelict, empower, people_nature, save_place, choose, people_enjoy, lucky, movement, improve, people_people, write, challenge, country, mind, young_people, talk, child, work, back_century	19
t_35	blanket_bog	national_park	good, fire, problem, landscape, people, management, time, policy, manage, create, understand, change, risk, possibly, positive, switch, low, major, general, deal, issue, connotation, direction, effect, amount, moment, biodiversity, bite, biomass, conservationist, negative, public, putting, drive, question, rewilding, set, end, hill, thing, manage_biomass, ngo, risk_management, economy, agricultural, intensive, worth, mistake, practice, act	5

t_36	south_pennines	national_park	landscape, pennines, south_pennines, south, industrial, dark, past, wilderness, tend, natural, put, area, people, activity, human, part, cover, reservoir, visit, talk, interest, bit, character, wild, word, thing, pennines_good, twohundred_year, twohundred, upland_landscape, gentle, scale, similar, add, modern, type, upland, history, area_cover, human_activity, infrastructure, manage_landscape, transform_landscape, visit_area, addition, derwent_valley, natural_landscape, derwent, actual, conservation	26
t_37	long_time	young_people	community, work, village, thing, people, time, instance, family, eyam, long, home, area, car, boy, shop, significant, bus, long_time, church, part, close, week, young, good, simply, property, rural, big, load, send, travel, husband, win, derbyshire, put, sort, community_good, thing_village, facility, rural_area, wait, council, drive, realise, lot, money, car_work, fortunate, horrible, husband_work	16

Supplementary material 3

Code book and short code for additional legend description for the Figure 5-4.

Code	Short code	Sub-code	Short sub-code	Manual Topic	TM Topic
Pro-environmental Action	PEA	Biodiversity	BIO	1	16
Pro-environmental Action	PEA	Blanket bog regeneration	BBR	2	11
Pro-environmental Action	PEA	Renaturalisation, rewilding, reforestation	REW	3	13
Pro-environmental Action	PEA	Traditional/historical farming (adoption of new methods - old knowledge revived for positive outcome, tradition as driver for sustainable coexistence)	TFA	4	26
Pro-environmental Action	PEA	Climate change emergency (working for awareness and nature protection)	CLIM	5	35
Pro-environmental Action	PEA	Heritage management/stakeholder engagement/activist groups	CHM	6	10
Challenges	CHA	Free access (free access for visitors, access to housing market for holiday homes, traffic)	ACS	7	7, 33
Challenges	CHA	Traditions and industries (grouse moors, quarries)	TRA	8	27

Change and Continuation	CAC	General perception	GEN	9	23
Change and Continuation	CAC	Diversification	DIV	10	5
Change and Continuation	CAC	Cement works and quarry (traditional works damaging landscape)	QUA	11	2, 20
Change and Continuation	CAC	Traditional farming (resistance to new methods, when tradition becomes the issue, when change is negative)	FAR	12	14
Change and Continuation	CAC	Heritage, intangible - connecting to landscapes through traditions, skills and traditional crafts	HIN	13	8, 19
Change and Continuation	CAC	Heritage, tangible - connecting to fabric, building material, object	HTA	14	32
Communities	COM	Community life - general engagement	ENG	15	21, 30
Communities	COM	Community life - personal engagement	ENG	16	37
People/Place Engagement	PPE	Connecting people and landscape - foster interest in the heritage landscape	INT	17	1, 12
People/Place Engagement	PPE	Connecting people and landscape - engage young and old "insiders"	INS	18	28
People/Place Engagement	PPE	Connecting people and landscape - engage young and old "outsiders"	OUT	19	3, 34

Place History	PLH	Connection through history - general local history	GLH	20	6, 17, 24
Place History	PLH	Connection through history - personal family history	PFH	21	22, 25
Place History	PLH	Connection through history - personal family history connected to local history	PFH	22	29
Landscape quality	LSQ	Connection through landscape quality - active use (walk, run, other activities, swimming)	ACT	23	4
Landscape quality	LSQ	Connection through landscape quality - appreciation of "wilderness"	WIL	24	18
Landscape quality	LSQ	Connection through landscape quality - appreciation of natural features	NAT	25	31
Landscape quality	LSQ	Connection through landscape quality - features of cultural landscapes/man-made heritage landscapes	CUL	26	15, 36
Landscape Quality	LSQ	Sense of belonging - genius loci effect	GEL	27	9

D. Appendices for Chapter 8:

Appendix 1a: Lab book

Introduction

Of the three methods used for data collection, the survey was found to be the most successful and useful to provide a sufficiently broad and deep insight into values of people living and working in the study areas. The method has the potential to be repeated on a regular or continuous basis, allowing a dynamic update of the *Social Landscape Characterisation* dataset.

The following lab book will give the code for creating and analysing a survey using *Qualtrics Survey* software with an embedded Google Map application, as shown in **Chapter 4**. The analysis methodology, using Natural Language Processing (NLP), Named Entity Recognition (NER) and Topic Modelling (TM) can also independently be applied to other datasets if these were collected using social media texts or interview transcripts, as show in **Chapter 3** and **5**. RStudio is used for the application of the developed code provided in the R programming language. QGIS was used for the spatial representation of the data. The steps in this methodology can be adapted or amended as appropriate.

The following will detail the application of the method from the data collection to the final visualisation of social values across the chosen project area. The lab book at the end of this methodology will enable this process (see also workflow diagram **Chapter 4**, Figure 4-1).

Survey design

The survey designed for this research, consisted of a three-part questionnaire and an embedded map to allow people to locate and pinpoint their favourite locations in commonly known and familiar format.

The questions allowed a wide range of information about the participants to be gathered. However, according to best practices in survey creation, only as much information should be collected as is needed to achieve the aim of the survey[1]. The questionnaire gives a range of questions that can be used to created tailored surveys for specific queries. For the questions used in this survey (**Chapter 4**) see **Appendix for Chapter 4/Supplementary Material 3**.

The Google map can be embedded in the Qualtrics survey, using HTML or JavaScript (provided in **Appendices for Chapter 8/2**). This allows participants to put a pin as location marker on the map and add stories in text form or/and upload images or photographs to illustrate the connection to their favourite places. The use of Google Maps has the advantage that location data from the markers come in latitude-longitude format, which can be directly uploaded to a GIS. The option to locate a favourite place giving the places name as text was included for people who were unable to navigate the map. However, the subsequent analysis and visualisation in GIS requires coordinates. If you intend to allow people to identify a location by text entry, the area-specific NER method (described below) can be used to add coordinates to the locations. The code is provided in the lab book, but a place-specific gazetteer has to be

---

[1] **Important note:** Before you start collecting personal data ensure that you have considered ethical and GDPR requirements and legal obligations for data protection. Make yourself familiar or seek advice on the legal obligation before dealing with personal data. A *Consent Form* and *Information Sheet* can be linked to the survey and should always be available for download to all participant (examples given in **D: Appendix for Chapter 8/3** and **4).**

compiled by filtering location data of datasets available for the particular area of interest (see **Chapter 3**).

NLP and data pre-processing

The provided code will pre-process the data which was provided by the *Qualtrics* survey in an Excel spreadsheet format. For the data cleaning and preparation, the data will go through a series of steps: extracting the required columns of the spreadsheet for this particular analysis, reordering values (coordinates) and renaming columns, conversion into a string (a sequence of characters), removal of special characters, stopwords, punctuation, and white space. The cleaned data is then converted into a document-text matrix (DTM), which can be fed into the TM code.

Analysing people's stories – Topic Modelling

The topic modelling code is executed on the DTM and calculates the themes latent within the documents on a statistical basis[2]. The parameters were set to 4,000 iterations of the modelling process in order to find the appropriate number of topics for the TM, which will depend on the dataset, the code will be run different models for k (number of topics). The modelling parameter is set to create models for 40 and 80 topics. In order to decide on the best number of topics, a graphical representation of topic coherence is provided. The point where the curve flattens out defines the optimal k number over the coherence of the themes. A smaller number would not give a fine-grained theme list and a greater number would create too fine-grained clusters.

The cluster diagram gives an impression of labels generated automatically by the algorithm based on the most frequent bi-grams (two closely associated words) within each cluster[3]. These labels resulting from the automated process can give an

[2] https://cran.r-project.org/web/packages/textmineR/vignettes/c_topic_modeling.html

[3] https://github.com/tqx94/Text-Analytics_LDA

idea of the overarching theme of a cluster. To assess the quality and coherence of clusters and label, a manual evaluation of the consistency and performance of the analysis is required in the next step. For this purpose the results up to this point are exported to a CSV file, which can be easily edited in a spreadsheet software.

Analysing people's stories – sentiment analysis

The sentiment analysis of the survey is based on the documents provided (text-based)[4]. Based on a training dataset the words within the documents will be tested if they match neutral, positive, or negative words. The sentiment will be given as a score between -1 and 1. Scores around towards 0 are seen to be neutral. The overall score of neutral, positive, and negative value represents the sentiment of each document. This score can be added to the csv as output at the end of the algorithm to visualise issue in the landscape at the stage of GIS visualisation.

Direct observation and categorisation

Direct observation and manual assessment of the result and performance of the algorithm is necessary to ensure that the documents are appropriately clustered and to refine the labels for the topics. This process can also help to understand the data in more detail. As this process is time-consuming, it should be a first step to process training data for the predictive modelling (this process lies beyond the scope of these guideline and is proposed for further research on the subject). The documents should be compared to each other in each cluster, to the topic label and to other similar clusters, which can be found in the cluster dendrogram. The cluster dendrogram represents cluster that have close similarities on nearby branches of the tree. The

[4] https://github.com/cjhutto/vaderSentiment

more topics (k) were chosen in the initial modelling phase the more related clusters will be created, less topics bunch clusters together. It now must be decided if clusters can be combined under one topic label. Topic label refinement means finding an overarching heading for the cluster that best represents the documents or stories within the cluster. These labels will then represent the categories and subcategories of the analysis. The manual labels need to be added to the spreadsheet as a new column.

Visualisation in GIS

The resulting analysis output can be created as csv and imported in QGIS. Maps created from the dataset can be classified based on the categories. The categorised format allows to visualise the categories based on the topic labels. Heatmaps or hotspot maps can be rendered from the symbol tab or with the heatmap plugin. Photographs are included in the file (attribute table) to create story maps with additional information. Useful is the webmap plugin for QGIS that allows to create interactive webmaps. This format can be integrated into websites and online publications[5].

[5] The QGIS User Guide is a helpful resource to create the outputs mentioned in this guide (https://docs.qgis.org/3.28/en/docs/user_manual/index.html)

Appendix 1b: Lab book with code for data analysis

Step 1: preprocessing

Load required packages/libraries

In this step all necessary R packages/libraries will be installed.

```
library(readxl)
library(stringr)
library(dplyr)
library(writexl)
library(csv)
```

Import data from Qualtrix Survey (Excel, csv)

The excel spreadsheet (csv file) generated by the Qualtrics survey (survey file) will be imported in for further processing.

```
datadf <- read_xlsx("./data_20220514.xlsx", skip = 1)
```

Select and rename columns

Not all columns of the survey file are necessary for the analysis (this depends on the data that will be analysed in the next steps). For the purpose of the sentiment and topic modelling analysis, the columns selected for all places are: locations (column 39 in survey file), places (column 40) stories, change and photographs (for each of the five places in columns 41-43, 48-50, etc.).

```
df <- datadf%>%
        select(39,(40:43),(47:50),(54:57),(61:64),(68:71))
```

The column headings were automatically created by Qualtrics based on the Questions. The selected columns will be renamed in the next step.

```
col_names = c("map_pin_location",
              "place_a",
              "place_a_story",
              "place_a_change",
              "place_a_photo_name",
              "place_b",
              "place_b_story",
              "place_b_change",
              "place_b_photo_name",
              "place_c",
              "place_c_story",
              "place_c_change",
              "place_c_photo_name",
              "place_d",
```

```r
              "place_d_story",
              "place_d_change",
              "place_d_photo_name",
              "place_e",
              "place_e_story",
              "place_e_change",
              "place_e_photo_name"
              )

names(df) <- col_names
```

Prepare data for further processing

This step prepares the coordinates for further processing in the next step. The output is a single vector of coordinate values (not pairs).

```r
# This function takes a string, replaces parentheses and spaces with c
ommas and returns a vector of numbers
cleanMapLocs <- function(text) {
  # removes the characters "(", ")", " "
  text = gsub("\\(|\\)| ","",text)
  # splits the remainder at each comma
  strVec = strsplit(text, ",")[[1]]
  # converts it into numbers (as.numeric) and
  # returns it as a tibble (table thingy)
  tibble(as.numeric(strVec))
}

#apply the function to every row in column map_pin_location
df$map_pin_location <- mapply(cleanMapLocs, df$map_pin_location)
```

This routine will go through the original dataframe and write out a new data frame that contains the data for each place that was selected by the participant in a separate row, i.e., if a participant selected 5 places, 5 rows will be created in the new data frame.

```r
# initialise our counter (this will count the number of places we foun
d and thus
# the rows in the new data frame)
i <- 0

# column name stems for the different places
labels = c("place_a", "place_b", "place_c", "place_d", "place_e")

# create an empty data frame
placesdf <- data.frame(Lon=double(), Lat=double(), Place=character(),
Story=character(), Change=character(), photo_name=character())

# loop over the original data frame row by row:
for (row in 1:nrow(df)) {

  # read the content of the map_pin_location field for each row
  # and put it into the variable "loc" (for easier readability)
```

```
  locs = df$map_pin_location[row][[1]] #[[]] indexing into the tibble

  # check if we do have some location data, or if the field is empty (
NA)
  # if it is not empty, then go into the inner loop, if it is empty,
  # the next bit (inner loop) is not executed and we go to the next ro
w (outer loop)
  if (!is.na(locs[1])) {

    # inner loop, we loop over the length of the "locs" vector in step
s of 2
    # (because each location has two numbers, lat and lon)
    # so this loop will be: j = 1, 3, 5, 7, 9 (if we have 5 pairs)
    for (j in seq(from=1, to=length(locs), by=2)) {

      # we increment our counter (this counts the number of locations
we got)
      i <- i+1

      # lat is the first number and lon in the second number in each p
air,
      # so for the each pair we have to get j and j+1 element from the
vector
      lat <- locs[j]
      lon <- locs[j+1]

      # for ease of numbering and to be less confusing, define n to be
the
      # number of the pair, so this will be:
      # n = 1, 2, 3, 4, 5
      n <- (j+1)/2

      # assign the right place name (place_a to place_e)
      placeName <- labels[n]

      # and fill the columns of the data frame with the data we want.
      # row is i
      # data to go into the columns are:
      # lat, lon, placeName, placeName_story, placeName_change, placeN
ame_photo_name
      placesdf[i,"Lat"] <- lat
      placesdf[i,"Lon"] <- lon
      placesdf[i,"Place"] <- df[row, placeName]
      placesdf[i,"Story"] <- df[row, paste(placeName, "story", sep="_"
)]
      placesdf[i,"Change"] <- df[row, paste(placeName, "change", sep="
_")]
      placesdf[i,"photo_name"] <- df[row, paste(placeName, "photo_name
", sep="_")]
    }
  }
}
```

Show all stories in the data frame

```
# how many stories do we have?
placesdf$Story[!is.na(placesdf$Story)]
```

show all photos in data frame

```
# how many photos do we have?
placesdf$photo_name[!is.na(placesdf$photo_name)]
```

Create a field identifier column (running number)

```
placesdf <- dplyr::mutate(placesdf, fid = row_number())
```

Export data

In the final step, export the places data frame to RDS, Excel and CSV:

```
saveRDS(placesdf, "placesdf.rda")
write_xlsx(placesdf, "placesdf.xlsx")
as.csv(placesdf, "placesdf.csv")
```

Step 2: data cleaning and DTM creation

Load required packages/libraries

In this step all necessary R packages/libraries will be installed.

```
library(dplyr)
library(readxl)
library(textmineR)

library(dplyr)
library(tidytext)
library(tidyr)
library(SnowballC)
library(textstem)
library(stringr)
library(tm)

library(wordcloud)
library(ggplot2)
library(RColorBrewer)
```

The file placesdf.csv that was exported in the last step needs to be uploaded into QGIS to delete points that are not within the study areas. The resulting dataset will be exported as placesdf.xlsx. The new dataset will only show data located within the study area.

Import data from QGIS exported file

The excel spreadsheet exported from the QGIS data cleaning step will be imported for further processing.

```
textdf <- read_xlsx("./placesdf.xlsx")
textdf <- textdf%>%
            select(Place, Story, Lon, Lat)
```

Create a field identifier column (running number)

This fid will allow to recreate stories from unigram lists in the following step.

```
textdf <- dplyr::mutate(textdf, fid = row_number())
```

Data cleaning of the stories

Delete all but alphabetical characters:

```
textdf <- textdf %>%
            mutate(cleantext = str_replace_all(Story, "[^[:alpha:]|']"
, " "))
```

Remove stop words, stem and lemmatise. Creates new data frames: unigram and unigram_cleaned, which contain single words (unigrams), their stems and lemmata, referenced by fid and Story.

```
stopw <- stop_words$word[stop_words$lexicon=="snowball"]%>%append("she
ffield")%>%append("peak")%>%append("district")
unigram <- textdf %>% unnest_tokens(word,cleantext, token="ngrams", n=
1)
 unigram_cleaned <- unigram %>% filter(!word %in%stopw)
 unigram_cleaned <- unigram_cleaned %>% mutate(word = str_remove_all(w
ord,"\\bna\\b"))
 unigram_cleaned <- unigram_cleaned %>% mutate(word = str_remove_all(w
ord,"\\b\\w?\\b"))
 unigram_cleaned <- unigram_cleaned %>% mutate(wordstem = wordStem(wor
d))
 unigram_cleaned <- unigram_cleaned %>% mutate(wordlem = lemmatize_wor
ds(word))
```

Regenerate stories (but using only the lemmas instead of the original words).

```
# sorts based on place and reconstructs story from lemmas for each pla
ce (combining stories for same place)
storydf <- unigram_cleaned %>% group_by(fid) %>% summarise(text = past
e(wordlem, collapse = " "))
# all stories in one string
storydf_nogroup <- unigram_cleaned %>% summarise(text = paste(wordlem,
collapse = " "))
#count all words in this one string
lengths(strsplit(storydf_nogroup$text[1], "\\W+"))

# create df with bigrams - useful? Not used.
df2 <- storydf %>% unnest_tokens(word, text, token = "ngrams", n=2)
```

Filter elements that do not have a story (empty).

```
library(rlang)
storydf_na <- storydf[(storydf$text==""),]
storydf <- storydf[!(storydf$text==""),]
```

Word frequency

The next step will create lists of most frequently used words and lemmas.

```
# wordcount frequency of normal words (cleaned no stopwords) but inclu
des walk and walking!!
word_count <- unigram_cleaned %>% group_by(word) %>% summarise(count =
n()) %>% arrange(desc(count)) %>% slice(1:10)
# wordcount frequency of lemma compared to the above
word_count_lem <- unigram_cleaned %>% group_by(wordlem) %>% summarise(
count = n()) %>% arrange(desc(count)) %>% slice(1:10)
```

Visualise data in graphs

The following step will create a barplot showing the most frequently used words in the stories.

```
ggplot(data = word_count) +
    geom_bar(aes(x=reorder(word, desc(count)), y = count), stat = "ide
ntity", fill = "#6699cc")+ theme_classic()+labs(x="words (as used)")
```

The graph created in this step will create a barplot of most frequently lemmatised words.

```
ggplot(data = word_count_lem) +
    geom_bar(aes(x=reorder(wordlem, desc(count)), y = count), stat = "
identity", fill = "#6699cc")+ theme_classic()+labs(x="words (lemmatize
d)")
```

Wordclouds can visualise the words used in stories showing their frequency by word size.

```
wordcloud(unigram_cleaned$word, max.words = 75, colors=brewer.pal(6,"D
ark2"))
```

The following will create a wordcloud from most frequently used lemmas.

```
wordcloud(unigram_cleaned$wordlem, max.words = 75, colors=brewer.pal(6
,"Dark2"))
```

Creating a DTM

Now we need to create the DTM for AI or Topic Modelling. For AI models that need training, two DTMs need to be created: training and prediction. For statistical TM there is no need for this split. The following will also save the output for further analysis.

```r
# Split Data into Training and Testing in R

train_split = 0.0 # proportion to be used in training data - select 0
for no split (statistical or no-training models)

if (train_split != 0){
sample_size <- floor(train_split*nrow(storydf))#1.0 dont separate
set.seed(777)

# randomly split data in r
picked <- sample(seq_len(nrow(storydf)),size = sample_size)
trainingdf <- storydf[picked,]
testdf <- storydf[-picked,]

dtm <- CreateDtm(trainingdf$text,
                 doc_names = trainingdf$fid,
                 ngram_window = c(1, 2))

testdtm <- CreateDtm(testdf$text,
                  doc_names = testdf$fid,
                  ngram_window = c(1, 2))

save(testdtm,file="testdtm.rda")
save(trainingdf,file="trainingdf.rda")
save(testdf,file="testdf.rda")
save(picked,file="picked.rda")
} else {
dtm <- CreateDtm(storydf$text,
                 doc_names = storydf$fid,
                 ngram_window = c(1, 2))
}

save(dtm,file="dtm.rda")
save(textdf,file="textdf.rda")
save(storydf,file="storydf.rda")
```

Step 3 - Topic Modelling

Load required packages/libraries

In this step all necessary R packages/libraries will be installed.

```
library(textmineR)
library(digest)
library(ggplot2)

library(tibble)
library(purrr)
library(dplyr)

library(reshape)

library(readxl)
library(writexl)
```

Load the objects saved in the last step from the RDA files.

```
load("dtm.rda")
load("textdf.rda")
load("storydf.rda")
```

Topic Modelling

Run the topic modelling process

```
# How many models to test?
# Model will be run with 1 to k_max topics to allow selection of the b
est number of topics to use
k_max = 60
k_list <- seq(1, k_max, by = 1)

# Create a model "name" based on the sha-hash of the DTM, this will on
ly change if the DTM has changed
# so we don't have to run this if the DTM is the same as before (this
step takes some time).
# This is used as the directory to save the models in.
model_dir <- paste0("models_", digest(dtm, algo = "sha1"))

# Only run these if the model directory dies not exist, else just read
the files.
if (!dir.exists(model_dir)) dir.create(model_dir)

  model_list <- TmParallelApply(X = k_list, FUN = function(k){
    filename = file.path(model_dir, paste0(k, "_topics.rda"))

  if (!file.exists(filename)) {

    # Fit a Latent Dirichlet Allocation model
```

```
    set.seed(12345)
    m <- FitLdaModel(dtm = dtm,
                     k = k,
                     iterations=2000, # I usually recommend at least 5
00 iterations or more
                     burnin = 300,
                     optimize_alpha = TRUE,
                     calc_likelihood = TRUE,
                     calc_coherence = TRUE,
                     #calc_r2 = TRUE,
                     cpus = 8)
    m$k <- k
    # Save to file
    save(m, file = filename)
   } else {
     load(filename)
   }
  m
})

#export=c("dtm", "model_dir")) # export only needed for Windows machin
es
```

Coherence score

The coherence of the documents in each cluster will be assessed in this step. The point where the curve flattens should be chosen for the optimal number of topics (k).

```
#model tuning
#choosing the best model based on the coherence

coherence_mat <- data.frame(k = sapply(model_list, function(x) nrow(x$
phi)),
                            coherence = sapply(model_list, function(x)
mean(x$coherence)),
                            stringsAsFactors = FALSE)
ggplot(coherence_mat, aes(x = k, y = coherence)) +
  geom_point() +
  geom_line(group = 1)+
  ggtitle("Best Topic by Coherence Score") + theme_gray()+#theme(panel
.grid.minor = element_blank()) +
  scale_x_continuous(breaks = seq(2,k_max,2)) + ylab("Coherence")

#ggsave("topicoherence.svg")
```

Select the optimal topic number

K is the number of topics that each model creates. The number is a choice of detail in the cluster against a too fine-grained approach. The more topics the better the coherence of the stories, but the less useful for a meaningful categorisation in general terms, meaning underfitting will not form proper clusters.

```
# select k_opt based on max coherence
k_opt <- which.max(coherence_mat$coherence)
# override by manually assigning k_opt
# k_opt = 38

model <- model_list[k_opt][[ 1 ]]
model$top_terms <- GetTopTerms(phi = model$phi, M = 20)
top20_wide <- as.data.frame(model$top_terms)
```

Likelyhood and R-squared

R-squared is the proportion of variability in the data as explained by a model. This means it evaluates how good the fit in topic models is. There is a risk of overfitting if R-squared increases (with estimated numbers of topics in LDA). Log-likelyhood compares different models and calculates a score for unseen documents and the learning rate: phi (words (token) over topics). Iterations through the process converge towards a maximum that flattens out. The optimal rate in our case is at approx. 500 iterations, after which the log-likelyhood presents a maximum and, therefore, a better model.

```
# R-squared
# - only works for probabilistic models like LDA and CTM

model$r2 <- CalcTopicModelR2(dtm = dtm,
                phi = model$phi,
                theta = model$theta)

model$r2

# log Likelihood (does not consider the prior)
#svg(file="iteration.svg", width=12, height=6)
plot(model$log_likelihood, type = "l")

#dev.off()
```

Visualising coherence and prevalence

Model coherence measures the score of each topic based on the semantic similarities between the words in each topic.

```
# probabilistic coherence, a measure of topic quality
# this measure can be used with any topic model, not just probabilisti
c ones
summary(model$coherence)

#svg(file="coherence.svg", width=12, height=10)
hist(model$coherence,
     col= "blue",
     main = "Histogram of probabilistic coherence")

#dev.off()
```

```
model$top_terms <- GetTopTerms(phi = model$phi, M = 50)
#head(t(model$top_terms))

write_xlsx(as.data.frame(model$top_terms), "topictopterms.xlsx")
```

Model prevalence is based on the Dirichlet priors for alpha (topic over documents). For example, a high level of alpha will assign more topics to each document. Prevalence measures the dominance of topics in each document. The closer the data points fit to a rising linear trend.

```
model$prevalence <- colSums(model$theta) / sum(model$theta) * 100

#svg(file="prevalence.svg", width=12, height=8)
plot(model$prevalence, model$alpha, xlab = "prevalence", ylab = "alpha
")

#dev.off()
```

Naive topic labelling

The naive topic labeling tool is based on probable bigrams in the documents. Two labels will be generated to get a better idea of the cluster content to aid later assessment and maual categorisation.

```
#
model$labels <- LabelTopics(assignments = model$theta > 0.05,
                            dtm = dtm,
                            M = 2)

head(model$labels,39)
write_xlsx(as.data.frame(model$labels), "modellabels.xlsx")

# put them together, with coherence into a summary table
model$summary <- data.frame(topic = rownames(model$phi),
                            label = model$labels,
                            coherence = round(model$coherence, 3),
                            prevalence = round(model$prevalence,3),
                            top_terms = apply(model$top_terms, 2, func
tion(x){
                              paste(x, collapse = ", ")
                            }),
                            stringsAsFactors = FALSE)
```

The following steps gives the topic numbers, labels, coherence and prevalence scores.

```
model$summary[ order(model$summary$prevalence, decreasing = TRUE) , ][
1:20 , ]

modelsummarydf <- model$summary[ order(model$summary$prevalence, decre
asing = TRUE) , ]
```

Cluster dendrogram

The cluster dendrogram provides a graphical output of the labels based on phi (words (token) over topics), showing the coherence of clusters like branches of a tree. The closer the branches the closer related are the cluster to each other. Similar clusters can be bunched together during the manual assessment if the documents have a high coherence.

```
model$topic_linguistic_dist <- CalcHellingerDist(model$phi)
model$hclust <- hclust(as.dist(model$topic_linguistic_dist), "ward.D")
model$hclust$labels <- paste(model$hclust$labels, model$labels[ , 1])


#png(file="clusterdendro.png", width=1200, height=1000)
#svg(file="clusterdendro.svg", width=12, height=10)
plot(model$hclust)

#dev.off()
```

The model has data in the background. For each story we take out the theta values and topic labels and merge them with the story database to get the information of the topic modelling in human readable format.

```
# Reads theta from model dataframe

theta <- as_tibble(model$theta)

# define a function to find the location of the
# maximum theta value (which will be the main topic)
findTopic <- function(thetaVector) {
  if (max(thetaVector)-min(thetaVector) != 0){
    paste("t_", as.character(which.max(thetaVector)),sep="")
  }else{
    "t_0"
  }
}

# and apply that to every row (the "1" means go through the df by row)
# each row will be handed to the function as a vector
theta$topic <- apply(theta, 1, findTopic)

# read the input df (storydf) which has the fid
#storydf <- read_xlsx("storydf.xlsx")

# and add the fid to the theta dataframe (take care that the two are c
onsistent!)
theta$fid <- storydf$fid
#theta$fid <- as.integer(model[["data"]]@Dimnames[[1]])

# optional, could also merge the whole theta frame:
# extract the topic and fid column from theta:
topics <- theta[, c("fid", "topic")]


# and finally, merge the two dfs based on fid
```

```
merged <- list(storydf, topics) %>% reduce(full_join, by='fid')

# optionally, leave thetas in the df by merging theta df as a whole
thetaMerged <- list(storydf, theta) %>% reduce(full_join, by='fid')

allMerged <- list(textdf, topics) %>% reduce(full_join, by='fid')
#create dataframe containing topics and topic_labels
topiclabels <- modelsummarydf[,c("topic","label.label_1","label.label_
2","top_terms")]
allMerged <- list(allMerged, topiclabels) %>% reduce(full_join, by='to
pic')
```

Write out all the data for manual processing and labelling.

```
#save model
save(model,file="LDAmodel.rda")

# save theta
save(theta, file= "theta.rda")
write_xlsx(theta, "theta.xlsx")
# save dataframes
save(merged, file="merged.rda")
write_xlsx(merged,"merged.xlsx")
save(allMerged, file="allMerged.rda")
write_xlsx(allMerged,"allMerged.xlsx")
save(thetaMerged, file="thetaMerged.rda")
write_xlsx(thetaMerged,"thetaMerged.xlsx")
```

Step 4 - Sentiment Analysis

Load required packages/libraries

In this step all necessary R packages/libraries will be installed.

```
library(readxl)
library(dplyr)
library(purrr)

library(ggplot2)
library(lexRankr)

library(readxl)
library(writexl)

# Reticulate allows integration of Python code in R workbook
library(reticulate)
```

Load the objects saved in the last step as excel spreadsheets and select only specific columns and rows.

```
allMerged <- read_xlsx(paste("allMerged.xlsx", sep = ""))
```

```
# select only fid and cleantext
sentimentdf <- allMerged %>%
            select(fid,cleantext)

# drop empty stories
sentimentdf <- sentimentdf[!is.na(sentimentdf$cleantext),]
```

__Phyton integration__

The sentiment analysis is done in Python, since Vader is implemented in Python, and the R port is not feature-complete. You need to have Python and pip installed.

```
import subprocess
import sys

def install(package):
    subprocess.check_call([sys.executable, "-m", "pip", "install", pac
kage])

install("vaderSentiment")

from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
analyzer = SentimentIntensityAnalyzer()
```

__Sentiment on uncleaned text__

```
vaderscores = [analyzer.polarity_scores(text) for text in r.sentimentd
f['cleantext']]
```

Handing sentiment data back to the R data frame

```
sentimentdf$vaderscores <- py$vaderscores

# function to extract the compound score from the vader score
vcomp <- function(x) x$compound

# new column for the compound score
sentimentdf$vadercompound <- mapply(vcomp, sentimentdf$vaderscores)
```

__Sentiment categorisation__

This function is applied to name sentiments with positive, neutral, negative for barplot

```
sent <- function(x){
  if (x>0.05){
    return ("Positive")
  }
  else if (x< -0.05){
    return ("Negative")
  }
  else {
    return ("Neutral")
```

```
  }
}

sentimentdf$vaderc <- mapply(sent, sentimentdf$vadercompound)

save(sentimentdf, file = paste("sentimentdf.rda"))
```

Merge the sentiment scores into the survey data set.

```
# optionally, leave thetas in the df by merging theta df as a whole
sentimentMerged <- list(allMerged, sentimentdf) %>% reduce(full_join,
by='fid')
```

Plot Vader sentiment

```
sentiStorycomp <- as.data.frame(sort(table(sentimentMerged$vaderc), de
creasing = T))
sentiStorycomp

c3 <-c("a","b","c")
f <- cbind(sentiStorycomp, c3)
ggplot(sentiStorycomp, aes(x=100*Freq/sum(Freq),y = reorder(Var1,Freq)
, fill=c3))+
   geom_bar(stat="identity")+
   guides(fill="none")+
   geom_col(color="black")+
   scale_fill_manual(values=c("c"="black","b"="darkgreen", "a"="orange
"))+
   xlim(0,90) +
     labs(
          x = "Percent",
          y = "Sentiment of Stories") +
     coord_flip()

 #ggsave("surveysenti.tiff", units="in", width=2.5, height=2, dpi=300,
compression = 'lzw')
```

Issue Identification

Identification of issues is unreliable if done on the sentiment of the whole text, since survey responses of issues often contain several positive statements and only a single sentence with a negative sentiment - the potential issue.

To identify these, the sentiment analysis is repeated on a sentence-by-sentence basis and all stories with one or more negative scores will be highlighted, even if they are overwhelmingly positive otherwise.

```
sentencesdf <- unnest_sentences(data.frame(doc_id = sentimentMerged$fi
d, text = sentimentMerged$Story, stringsAsFactors = FALSE), sent, text
)

# drop empty sentences
sentencesdf <- sentencesdf[!is.na(sentencesdf$sent),]
```

```
vaderscores = [analyzer.polarity_scores(text) for text in r.sentencesd
f['sent']]

sentencesdf$vaderscores <- py$vaderscores

vcomp <- function(x) x$compound

sentencesdf$vadercompound <- mapply(vcomp, sentencesdf$vaderscores)

sentencesdf$vaderc <- mapply(sent, sentencesdf$vadercompound)

sentiSentcomp <- as.data.frame(sort(table(sentencesdf$vaderc), decreas
ing = T))
sentiSentcomp
```

Plot sentiment on sentences

```
c3 <-c("a","b","c")
f <- cbind(sentiSentcomp, c3)
ggplot(sentiSentcomp, aes(x=100*Freq/sum(Freq),y = reorder(Var1,Freq),
fill=c3))+
   geom_bar(stat="identity")+
   guides(fill="none")+
   geom_col(color="black")+
   scale_fill_manual(values=c("c"="black","b"="darkgreen", "a"="orange
"))+
   xlim(0,90) +
     labs(
          x = "Percent",
          y = "Sentiment of Sentences") +
     coord_flip()

 #ggsave("sentencessenti.tiff", units="in", width=2.5, height=2, dpi=3
00, compression = 'lzw')
```

Output of sentiment scores

```
write_xlsx(sentencesdf, "sentencesdf.xlsx")
```

Prepare the data set of issues (all negative sentences)

```
negativedf <- sentencesdf%>%filter(vaderc == "Negative")%>%group_by(do
c_id)%>%summarise(text=paste(sent, collapse = " "),vaderc = vaderc[1])
%>%rename("fid"="doc_id")%>%rename("Possible Issue"="text")%>%select(f
id, "Possible Issue")
```

Export

```
sentimentMerged <- list(sentimentMerged, negativedf) %>% reduce(full_j
oin, by='fid')
write_xlsx(sentimentMerged, "sentimentMerged.xlsx")
```

Appendix 2: Javascript for embedding a Google Map in a Qualtrics survey

```
// Needs a Google API registration. This is free for researchers. Incl
ude:
// Maps API
// ??
var googleMapAPIKey = " "; // enter your access code here

// Variable definitions, do not change
var allMarkers = [];
var dataBox;

Qualtrics.SurveyEngine.addOnload(function() {

    // --- User Variables, set these to match your map section ---
    // Centre of map, use geographic latitude and longitude in degrees
:
    var mapCenterLat = 53.38;
    var mapCenterLng = -1.58;

    var mapZoom =10;
    var mapWidth = "100%";
    var mapHeight = "500px";
    var locationInputWidth = "96%";
    var locationInputMargin = "2%";
    var locationInputPadding = "15px";

    // max. number of locations to be selected
    var maxMarkers = 5;

    // This is for the marker labelling on the map  A B C ...
    // add to the labels string, if more locations are needed!
    var labels = "ABCDEFG";
    var labelIndex = 0;

    // Get the data entry box and store it in a variable
    dataBox = document.getElementById("QR~" + this.questionId);

    // Get the question container and store it in a variable.
    var questionContainer = this.getQuestionContainer();

    try {
        // Create a map object and append it to the question container
.
        var mapObject = document.createElement('div');
        mapObject.setAttribute("id", this.questionId + "-map");
        mapObject.style.width = mapWidth;
        mapObject.style.height = mapHeight;
        questionContainer.appendChild(mapObject);
        var mapID = this.questionId + "-map";
    } catch (err) {
        console.log("Unable to create map object. Details: " + err);
        alert("An error occurred creating the map.");
```

```
    }

    // Hide the data box
    try {
        dataBox.style.display = 'none';
    } catch (err) {
        console.log("Unable to hide data box.");
    }


    // This function calls itself once per second until the Google Map
s API is loaded, then it displays the map.
    function displayMap() {
        try {

            var map = new google.maps.Map(document.getElementById(mapI
D), {
                center: {
                    lat: mapCenterLat,
                    lng: mapCenterLng
                },
                streetViewControl: false,
                zoom: mapZoom
            });


            // this bit controls the pins and text prompt to label the
pins
            var counterss = 0;
            google.maps.event.addListener(map, 'click', function(event
) {
                if (counterss > (maxMarkers - 1)) {
                    alert("To many markers on the map!") // stop from
having too many markers on the map
                } else {

                    counterss++ // iterates through all the markers

                   // var labeltext = prompt("Enter Marker Label Info"
);
                    var temp_marker = addMarker(event.latLng, map);

                    google.maps.event.addListener(temp_marker, 'dragen
d', function(event) {
                        writeData();
                    });
                    allMarkers.push(temp_marker);

                    // Write data to Qualtrics question
                    writeData();

                    var card = new map.card();
                    card.getBody().innerHTML = labeltext;
```

```
                    var labelcontent = JSON.parse(localStorage.getItem
('map'));
                    console.log(labelcontent);
                    if (a == []) {
                        var index = 0;
                    } else {
                        var index = a.length;
                    }
                }
            });

            function addMarker(location, map) {
                var marker = new google.maps.Marker({
                    position: location,
                    label: labels[labelIndex++ % labels.length],
                    map: map,
                    draggable: true
                });
                //attachNote(marker, note);
                return marker;
            }

        } catch (err) {
            setTimeout(function() {
                displayMap()
            }, 1000)
        }
        return allMarkers;

    }
    displayMap();
});

function attachNote(marker, note) {
    var infowindow = new google.maps.InfoWindow({
        content: note
    });
    marker.addListener('click', function() {
        infowindow.open(marker.get('map'), marker);
    });
}

// Load the Google Maps API if it is not already loaded.
try {
    if (typeof googleMapJS == 'undefined') {
        var googleMapJS;
        if (googleMapJS == null) {
            googleMapJS = document.createElement('script');
            if (googleMapAPIKey == "Your key" || googleMapAPIKey == nu
ll) {
                googleMapJS.src = 'https://maps.googleapis.com/maps/ap
i/js' + "?libraries=places";
            } else {
```

```
                googleMapJS.src = 'https://maps.googleapis.com/maps/ap
i/js?libraries=places&key=' + googleMapAPIKey;
            }
            document.head.appendChild(googleMapJS);
        }
    } else {
        console.log("Map already loaded.")
    }
} catch (err) {
    console.log("Unable to load Google Maps API. Details: " + err);
    alert("Unable to load Google Maps API.");
}


// Function writes the data to Qualtrix
function writeData() {
    dataBox.value = "";
    allMarkers.map(e => {
        dataBox.value += e.getPosition() + ", "
    });
};


Qualtrics.SurveyEngine.addOnUnload(function() {

});
```

Appendix 3: Consent form for online questionnaire

Informed Consent Form for survey participants

Beyond Landscape: Extending Historic Landscape Characterisation to develop a more inclusive and transparent approach to managing local heritage.

By taking part in this survey to consent to the following:

1. I consent to participate in this survey.
2. I have read and understood the information provided in the information sheet link.
3. I have understood that no identifiable personal information will be collected/responses will be anonymised.
4. I have understood that responses will be stored securely.
5. I have understood that the data resulting from this survey will be used in the project and secondary/further analysis.
6. I have understood that I can withdraw my response before the submission of this thesis.
7. I have understood that the result of this survey will only be used as detailed in the information sheet and will be published.
8. If you are uncomfortable to answer questions you can skip these.

Thank you very much for taking part in this survey and helping to advance this research. By ticking the box below, you will consent and can start the survey. If you have any other questions regarding this study, please contact the researcher (mt1451@york.ac.uk) or the academic supervisor John Schofield (john.schofield@york.ac.uk).

I consent to the details set out above. Yes ☐ No ☐

[Participants will be directed to the survey if indicated 'Yes' and directed to a 'Thank you' message skipping the survey when choosing the answer 'No'.]

Appendix 4: Information Sheet for online questionnaire

Participant Information Sheet

People and Places: Social Landscape Characterisation for inclusive and sustainable heritage management.

The University of York would like to invite you to take part in the following research project: *People and Places: Social Landscape Characterisation for inclusive and sustainable heritage management.*

Before agreeing to take part, please read this information sheet carefully and let us know if anything is unclear or you would like further information.

Background:
This study forms part of a PhD research project at the University of York. The information the participants of the online survey and in-depth interviews provide will give an insight into how communities perceive the character and what they value in the place in which they live. The world around us is constantly changing and developing and Historic Landscape Characterisation (HLC) was developed to assist heritage professionals and town planners to preserve the distinct character – what makes places special. The study will analyse how the views of heritage professionals and laypeople compare and how HLC can be enhanced to reflect the results of this study by including the values of communities into local heritage management.

Researcher:
The researcher is Martina Tenzer, an AHRC funded WRoCAH PhD researcher at the University of York, Department of Archaeology/Cultural Heritage Management (email: mt1451@york.ac.uk). The project is supervised by Professor John Schofield (john.schofield@york.ac.uk) and Professor Julian D Richards (julian.richards@york.ac.uk).

This research has been subject to ethical review by the University of York Department of Archaeology Ethics Committee. If you have any questions regarding the ethics process, please contact the Archaeology Ethics Committee member James Taylor (james.taylor@york.ac.uk).

Method:
Interviews will be carried out in spring 2022, the specific date and time will be arranged between the participant and the researcher. The interviews will be held via Zoom. The interview will last approximately 30–45 minutes and recorded. Following the interview, the researcher will prepare a digital transcript which can be provided to the interviewee if requested.

The online survey will run from September 2021 to May 2022 and provide information from heritage practitioners across England, and communities and individuals in Sheffield and the Peak District National Park. The aim is to build up a picture of the use of HLC in local authorities and the perception of heritage in local communities. By completing this survey, you will contribute to a database that holds the values of local people in local places. This will provide a vital background for local authorities.

Confidentiality, anonymity, security:

Your response will be recorded during the interview. To minimise the risk of loss or accidental sharing, the data will be immediately transcribed, anonymised and stored on a password-secured university server. The data will be held for 10 years after submission of the thesis and then securely destroyed. The measures for anonymizing and securing the data will minimise any identified risks for participants as follows:

- Loss and accidental sharing of data before or during transfer of data - Voicing an opinion that deviates from the official line of the local authority or the community
- Potentially emotionally charged topics concerning change/identity/planning issues

Why have I been invited to take part?

[Depending on the survey/group only the relevant passage will be included]

You have been invited to take part because you are a member of ALGAO and a potential user of HLC.

You have been invited to take part because you are a member of the Polish/German community living in Sheffield/the Peak District National Park and your opinion on the character of the place you live is important. You have been invited to take part because you are a resident of Sheffield/the Peak District National Park and your opinion on the character of the place you live is important.

Do I have to take part?

No, participation is optional. If you do decide to take part, you will be given a copy of this information sheet for your records and will be asked to complete a participant information form. If you change your mind at any point during the study, you will be able to withdraw your participation without having to provide a reason up until the submission of the thesis in January 2024.

On what basis will you process my data?

Under the General Data Protection Regulation (GDPR), the University has to identify a legal basis for processing personal data and, where appropriate, an additional condition for processing special category data.

In line with our charter which states that we advance learning and knowledge by teaching and research, the University processes personal data for research purposes under Article 6 (1) (e) of the GDPR:

Processing is necessary for the performance of a task carried out in the public interest

Special category data is processed under Article 9 (2) (j):

Processing is necessary for archiving purposes in the public interest, or scientific and historical research purposes or statistical purposes

Research will only be undertaken where ethical approval has been obtained, where there is a clear public interest and where appropriate safeguards have been put in place to protect data.

In line with ethical expectations and in order to comply with common law duty of confidentiality, we will seek your consent to participate where appropriate. This consent will not, however, be our legal basis for processing your data under the GDPR.

*How will you use my data***?**

Data will be processed and analysed as part of the Doctoral thesis, which will be marked by internal and external examiners. Staff and students of the University of York will have access to an electronic copy of this thesis that will be desposited with the White Rose eThesis Repository at the University of York. Access for researchers outside the university will be granted on request. Transcribed data may be used as secondary data as part of a further study. No access will be given to the original recording of interviews. The participant data will also be used for public talks at Conferences and events.

Further consent will be sought if the researcher wants to use the data of this project for any other form of public presentation.

Will you share my data with 3rd parties?

Yes. The following third parties will have access to your data, in the form of a report distributed across ALGAO members.
The anonymised data of interviews will be published as a journal article and the data of general online survey will be part of a web map accessible through the websites of the partner organisations (Sheffield City Council, SYAS, Peak District National Park Authority).

Anonymised data may be reused by the research team or other third parties for secondary research purposes.

Will you transfer my data internationally?

Possibly. The University's cloud storage solution is provided by Google which means that data can be located at any of Google's globally spread data centres. The University has data protection complaint arrangements in place with this provider. For further information see, https://www.york.ac.uk/it services/google/policy/privacy/.

What rights do I have in relation to my data?

Under the GDPR, you have a general right of access to your data, a right to rectification, erasure, restriction, objection or portability. You also have a right to withdrawal. Please note, not all rights apply where data is processed purely for research purposes. For further information see, https://www.york.ac.uk/records management/generaldataprotectionregulation/individualsrights/.

Questions or concerns

If you have any questions about this participant information sheet or concerns about how your data is being processed, please contact Jonathan Finch, Chair of the Arts and Humanities Ethics Committee (jonathan.finch@york.ac.uk) in the first instance. If you are still dissatisfied, please contact the University's Acting Data Protection Officer at dataprotection@york.ac.uk.

Right to complain

If you are unhappy with the way in which the University has handled your personal data, you have a right to complain to the Information Commissioner's Office. For information on reporting a concern to the Information Commissioner's Office, see www.ico.org.uk/concerns.

Appendix 5: Consent form for interviews

Informed Consent Form for interview participants

Beyond Landscape: Extending Historic Landscape Characterisation to develop a more inclusive and transparent approach to managing local heritage.

Thank you for taking part in this study of the University of York. Please read the accompanying information sheet and check the relevant boxes to indicate your consent regarding your participation and the use of your data.

I have read and understood the information sheet.	Yes ☐	No ☐
I consent to participate in this research project.	Yes ☐	No ☐
I have had opportunities to ask questions about the study.	Yes ☐	No ☐
I permit saving and using my responses anonymously.	Yes ☐	No ☐
I permit that my responses will be used anonymously for this thesis and for further/secondary analysis and research purposes.	Yes ☐	No ☐
I give consent to record the interview.	Yes ☐	No ☐
I have understood that I can withdraw from the study before the submission of this thesis and have been informed how to do this.	Yes ☐	No ☐

Please sign below (by typing your name) to indicate your consent.

Participant: Name and Date

Researcher

You can withdraw your consent until 1 January 2024. If you have any further questions regarding this study, please contact the researcher (mt1451@york.ac.uk) or the academic supervisor John Schofield (john.schofield@york.ac.uk).

Appendix 6: Ethical Approval

Name of Applicant: Martina Tenzer

Email Address: mt1451@york.ac.uk

Is this a collaboration with another researcher? No

Name of Additional Applicant:
Department Centre or Unit:

Are There Additional Researchers?
Name of Additional Applicant:
Staff/Student Status: PhD Student
Department Centre or Unit:

Are There Additional Researchers?
Please give the names, email addresses, and affiliations of any other researchers that need ethics approval for this project:

Staff/Student Status: PhD Student
Name of 1st Supervisor: Professor John Schofield
Email address: julian.richards@york.ac.uk

Name of 2nd Supervisor: Professor Julian D Richards
Email address: julian.richards@york.ac.uk

Title of Project: People and Places: Social Landscape Characterisation for inclusive and sustainable heritage management
Project Start Date: 2021-05-01
Duration: 2 years 7 months

Is this research under the jurisdiction of any other external ethics board? (e.g. the European commission; Human Subjects Review in the USA): No

Funded: Yes
Funding Source?: Arts and Humanities Research Council

Please briefly outline the questions or hypotheses that will be examined in the research. This can normally be copied from your research proposal.:
The project will develop a Social Landscape Characterisation related to the currently used Historic Landscape Characterisation as a background for planning decisions. The aim is to create a more transparent and inclusive decision-making process.

Research questions are:

1. How can social values or "soft", qualitative data - representing local knowledge and expertise - be identified, collected, analysed, and translated into visual representations?
2. What are the most suitable methods and tools for collecting and analysing of qualitative data to produce guidelines for a reliable and reproducible methodology?
3. How can a participatory approach benefit a community's well-being, economic and ecological resilience, and create transparent and inclusive decision-making through influence on local policies and legislation?

Methods in this part of the project are detailed in the attached Method Statement and include:

(1) Practitioner survey: evaluate the current use of Historic Landscape Characterisation (HLC) within local authorities and how communal values might be used if available. (2) Resident survey: explore social/communal values of individuals and communities. (3) Interviews with practitioners will deepen the understanding of attitudes toward the integration of social values in the planning process. (4) Interviews with members of the public will provide a deeper understanding of the definition of ‘everyday heritage’ and how much change is acceptable. (5) Walk-and-talk interview/observation: map and record the experience of everyday heritage during walks with residents (this will only take place in summer 2022 if the government legislation and university regulations permit it). (6) Case Study: A hypothesis tested amongst Polish and German nationals will explore if Polish nationals form stronger communities with a more visible identity (Polish bakeries and shops, Polish Catholic Church, and community centres) compared to the German community that is almost invisible regarding identity within the British society. (7) Pilot projects: Two small scale pilot projects – with heritage practitioners at my partner organisations and a group of volunteers from the National Park – to evaluate the usefulness and appropriateness of the questionnaires and to reflect on the approach preceding the large-scale fieldwork.

Additional method for the side project “Place attachment tweets”: Twitter data will allow me to understand how people's perception of their environment changed from before the Covid-19 pandemic and one year on from the first lockdown in the UK. The data will give me information on what places people are talking about, if attitudes to places changed, and how often people tweet about special places, buildings or landscapes. This additional source will inform my research compared to the pre-defined survey questions.
The initial collection of tweets will be carried out on a weekly basis over one month until sufficient data is collected. It is anticipated that this initial stage of data mining will provide me with several thousand individual tweets. General search for this part of the project will be based on hashtags, for example, #Sheffield, #Sheffieldissuper, #Kelham, #PadleyGorge, #Walkley, #madeinsheffield , #PeakDistrict, #MayDay. A Python programme will be used to clean the data on the basis of a keyword search which will be conducted using general keywords, such as ‘place’, ‘park’, ‘street’, ‘yard’, ‘walk’, ‘view’, ‘museum’, ‘hall’, ‘works’, ‘bike ride’, ‘subway’, ‘Railway’, ‘stone circle’, ‘brook’,

‘outside’, ‘edge’. These categories will be published in a similar generalised overview and categorisation of place attachment features. A sentiment analysis will be based on keywords compiled in a wordlist, containing items such as ‘miss’, ‘fun’, ‘sunset’, ‘memories’ , ‘alright’, ‘remember’, ‘wonderful’, ‘eerie’, ‘beautiful’, ‘enjoying’, ‘adorable’, ‘lovely’, ‘stunning’, ‘good’, ‘sad’. In addition, tweets containing photographs will be manually analysed.

Methods of data collection:

Online surveys, Zoom interviews, walk-and-talk interviews/participant observation (this will take place in summer 2022 only if government legislation and university regulations permit it). Twitter keyword, sentiment and longitudinal analysis (Twitter data will be collected at different times throughout the project to analyse tweets from three to four different points in time starting with data from the summer 2019 (before Covid-19), summer 2020, summer 2021 (after the lockdown series) and summer 2022).

How many participants will take part in the research?:
(1) (2) Survey Practitioners/public 50-80, (3) Interview Practitioners: 8-10, (4) (5) Interview public: 8-10, (6) Case Study Interviews: 20, (7) Pilot projects: each 5-8. Twitter research: up to 10 m tweets can be pulled/month to gain insight into the subject

How will they be invited to take part in the study?:
(1) (3) I will contact the Association of Local Government Archaeological Officers (ALGAO) secretary, introduce my project and ask to distribute the survey amongst the members in local authorities, participants will be asked if they are interested in a follow-up interview. (2) (4) (5) (7) I will also ask my partners at the Peak District National Park Authority and the Sheffield City Council/South Yorkshire Archaeological Services to publish the survey on their websites. Furthermore, the Women’s Institute, neighbourhood community groups (e.g. Walkley Community group, Friends of the Peak District, Bamford Community group) will be contacted by email and asked to distribute the survey amongst their members. These groups also use social media channels such as Facebook and Twitter that will be used to publish the survey after permission is granted by the group administrators. Interested participants will be invited to a follow-up interview and a walk-and-talk interview/observation according to regulations at that time. (6) I will also contact the group leaders of the Polish Catholic Church and community group, the owner of the Polish Deli Sheffield, the group facilitators of the German ‘Stammtisch’ (round table) and the German ‘Frauenstammtisch’ (Women round table). Potential participants will be invited by email describing the aim of the project and setting out how the interview will be conducted. Information sheets and consent forms will be distributed together with the invitations.

Confirm that you will obtain confirmed consent before subjects participate in the study: I will provide consent sheets for subjects to sign before participating in the study, I will retain these forms for the duration of the research. It is not possible and not anticipated to obtain consent from individuals for the Twitter research. However, tweets in this project do not contain sensible or confidential data (Twitter tweets are in general shared publicly). No individual tweets or personal data will be collected. Keyword and sentiment analysis will be applied to several thousand expected tweets to produce statistics which will not allow for individuals to be traceable. Keywords will be used in the form of general overviews, categories or paraphrased. Data mining on Twitter will be repeated over a period of weeks until an acceptable threshold of data for analysis is reached.Should the overall number of tweets be insufficient to allow for implicit anonymisation the twitter data will be deleted and not used in the project. (https://www.york.ac.uk/staff/research/governance/research-policies/social-media-data-use-research/) (https://www.york.ac.uk/staff/research/governance/research-policies/social-media-data-use-research/social-media-examples/)

Please upload your project information sheet to be given to all participants.: Tenzer_InformationSheet - Martina Tenzer.pdf
Please upload your informed consent form.: Tenzer_Consent_Form - Martina Tenzer.pdf
Please upload any additional files.: Tenzer_Method_statement - Martina Tenzer.pdf, Tenzer_1_7_practitioner_questionnaire - Martina Tenzer.pdf, Tenzer_2_7_public_questionnaire - Martina Tenzer.pdf

Are the results to be disseminated to the participants? : Yes

How will you be disseminating your results to your participants?:
ALGAO and members of the local authorities will receive a written report summarising the results. It is also anticipated to give a talk in the ALGAO Forum. Members of the public will have access to a web map that contains the information of the online survey accessible through the websites of the organisational partners (PDNPA and SYAS). The case study will be published in a journal article provided to the participants of the case study interviews. If indicated, the participants will get a pdf transcript of their interview and access to the PhD thesis after completion, which will not be publicly available. According to the Twitter policy, publishing of aggregated data and stats are permitted. The Academic Twitter Developer account has been approved by Twitter. This approval also allows that aggregated data - but no primary, raw or individual data - will also be shared with local government organisations (Sheffield City Council and Peak District National Park Authority).

Are you ensuring anonymity for your participants?:
Yes

Please explain how you plan to anonymise data or pseudonymise data during the project to minimise data protection risk. The identity of the participants remains anonymous, and all reasonable efforts will be undertaken to ensure that participant information and personal details cannot be linked to participants. Survey questionnaires will not ask for personal information such as name, date of birth, address. Information about gender, age, nationality will be used for statistical purposes only. Interviews will be anonymised or pseudonymised. Complete anonymity cannot be guaranteed for members of local authorities or in focus groups with very small group

sizes. This will be explained to the participants, and anonymity will be ensured as much as possible in the circumstances.
Data from Twitter research will be aggregated analysis and statistics. No individual tweets, comments or statements will be published in my research. This approach is conforming to policies and agreements of the Twitter Developer platform. An academic account application for the Twitter Developer account has been submitted in advance of this application and approved by Twitter Inc. on 19 April 2021. Aggregated data and stats resulting from the analysis will not allow linking back to individual tweets or individuals through online search. Location data will conform to Twitter policies and only consist of "heat maps and related tools that show aggregated geo activity" (https://developer.twitter.com/en/developer-terms/policy). Also, location data will focus on the places mentioned in the tweets which is directly connected to the tweet content not to the geolocation of the individual (https://developer.twitter.com/en/developer-terms/agreement).
To avoid re-identification of the Twitter users analysis of Tweets will be mainly in the form of statistics or agglomerated data that will use category keywords, e.g. 'park', 'canal', 'views', 'landscape', 'item in the Sheffield HER dataset', 'assets contained in monument data', 'traditions' instead of the words used in the original tweets. Only if a number of individual tweets of a specific place is reached that would not allow a re-indentification of an individual person the place will be defined more precisely., e.g. more than 30 people in Sheffield mentioned 'Greaves Park' in a tweet. Tweet content will be reworded or summarised paraphrasing the content to avoid identification of individual tweets.(Ahmed, W., Bath, P. and Demartini, G. (2017) Chapter 4 Using Twitter as a Data Source: An Overview of Ethical, Legal, and Methodological Challenges. In: Woodfield, K., (ed.) The Ethics of Online Research. Advances in Research Ethics and Integrity (2). Emerald , pp. 79-107.)

If anonymity is not being offered please explain why this is the case.:
Anonymity is being offered.

Please explain the measures in place to ensure that you are capturing the minimum amount of personal data/special category data necessary for your research project. The questionnaires will be discussed with my supervisors. Two pilot projects will be undertaken in advance to test the questions, and feedback from this first stage will feed into a revision of the survey questions. This reflection will ensure that the data collected is useful and will enhance my project. Data such as age, gender and nationality are necessary for statistical purposes; however, this statistical data will not be directly linked to individuals participating in the surveys or interviews.

Please detail the types of data you will be collecting.: Interviews, Questionnaires, Audio recordings, Video recordings, Photographs, creative essays, mapping of walks. Analysis of Twitter data.

Where will the data be stored electronically?: Password protected PC, GOOGLE drive with no sharing enabled

Where is the data to be stored in paper form?: Locked filing cabinet

At what point are you proposing to destroy the data, in relation to the duration of this project?: Ten years after the research is completed

How will you destroy this data?: Secure delete it electronically, Shred the paper forms

If you are sharing your data with others outside your department, what steps are you taking to ensure that it is protected?: Sharing via password protected Google Drive

If you are sharing personal or special category personal data with others outside your department, what steps are you taking to ensure that it is protected? If you are working collaboratively with third parties or sharing data with non-University personnel, please ensure that you have consulted the Information Governance Office and/or IP and Legal to ensure appropriate contracts and/or data sharing arrangements are in place.:
No personal data will be shared. Only statistical analysis will be shared with others.

Are you exporting this data outside the EU?: I am not exporting it outside the EU

If the data is to be exported outside the European Union, what steps are you taking to ensure that it is protected? Note: you must identify how you will comply with General Data Protection Regulation requirements. https://www.york.ac.uk/records-management/dp/:

Risks to participants (e.g. emotional distress, financial disclosure, physical harm, transfer of personal data, sensitive organisational information). All research involving human participants can have adverse effects. The answer of "none" will not be accepted.
Risks to participants:

- Recording of the interview for transcription, transferring of personal data during the online data collection process and providing an email to express interest in participation in follow-up interviews.
- Expressing a personal opinion that may not conform with the current policies in local authorities causing distress and frustration.
- Expressing a professional opinion that diverges from current planning policies and regulations deployed by the planning authority causing an inner conflict for the participant.
- Discussing the topic of identity with Polish and German migrants can cause emotional distress during the interview.
- Harm from traffic during walk-and-talk interviews.
- Publishing data from social media that can be traced back to individuals without consent.

- Publishing personal opinions from social media content.

Please state how you will mitigate these risks to participants.: Data will only be collected with expressed consent and only be used for the purpose detailed in the information sheet. The data will be anonymised, securely stored, and securely destroyed after 10 years. Participants can receive a copy of their responses and withdraw from the project before submission.
Participants will be anonymised, where possible persons should not be identifiable through their responses. Participants will be made aware that they can withdraw their data from the project before submission.
Questions in interviews and surveys may be skipped if participants find these controversial. Participants will be anonymised and information generalised and summarised where possible. Participants can receive a response sheet and withdraw before submission of the thesis. The interviewer will avoid questions that can be controversial or politically sensible.
The interviewer will avoid asking questions that can cause distress and clarify at the beginning of the interview that the interviewee does not have to answer questions that would cause irritation or frustration.
All necessary caution will be taken when undertaking walk-and-talk interviews with participants regarding traffic.
All data consulted for the Twitter research is 'public', nevertheless, no individual comments or content of tweets will be published. Twitter analysis will focus on keyword and sentiment analysis. Only analysis results and stats will be stored and published.

Risks to researchers (e.g. personal safety, physical harm, emotional distress, risk to accusation of harm/impropriety, conflict of interest....) The answer of "none" will not be accepted.:
Risks to researcher:

- Risk of loss or accusation of mishandling data.
- Harm during walk-and-talk interviews (subject to government policies and university regulations in 2022).

Please state how you will mitigate these risks to the researcher.:
Data will be collected, stored, and used in accordance with the ethics procedures detailed in this document.
Care will be taken regarding traffic and the public during the walk-and-talk interview.

University/institutional risks (e.g. adverse publicity, financial loss, data protection....) The answer of "none" will not be accepted.:
The researcher is a representative of the university and research should be conducted as unbiased as possible. There is still a risk that the researcher may be accused of following a personal political agenda or misleading interviewees with biased, 'loaded' questions.

Please state you will mitigate these risks to the university. The sampling will be random, and participation voluntary. The interviewer/researcher will reflect on feedback from pilot projects to avoid personal unconscious bias. The research will not voice personal experience or political opinions during the interviews.

Financial conflicts of interest (perceived or actual with respect to direct payments, research funding, indirect sponsorship, board or organisational memberships, past associations, future potential benefits, other....):
None

Please draw our attention to any other specific ethical issues this study raises.:
None

Please tick if true, otherwise leave blank::
Informed consent will be sought from all research participants, All data will be treated as anonymously as possible and stored in a secure place, All relevant issues relating to General Data Protection Regulation have been considered (see https://www.york.ac.uk/records-management/generaldataprotectionregulation/) &, if necessary, the Data Protection office contacted (Dr Charles Fonge, Borthwick Institute, charles.fonge@york.ac.uk), All quotes and other material obtained from participants will be anonymised in all reports/publications arising from the study where appropriate, All reasonable steps have been taken to minimise risk of physical/psychological harm to project participants, All reasonable steps have been taken to minimise risk of physical/mental harm to researchers, Participants have been made aware of and consent to all potential future uses of the research and data, Any relevant issues relating to intellectual property have been considered (see https://www.york.ac.uk/staff/research/external-funding/ip/policy/), There are no known conflicts of interest with respect to finance/funding, The research is approved by the Supervisor, Head of Department or Head of Research

Please explain in the space below, why if any of the above items have not yet been confirmed::
N/A

Are there any issues that you wish to draw to the Committee's attention? It is your responsibility to highlight any ethical issues that may be of perceived or actual interest. : None

Type your name to sign the document: Martina Tenzer

Appendix 7: Ethical Approval – updated method statement

Updated method statement to facilitate the developments in the analysis approach using Artificial Intelligence

provided to the ethics committee on 03/01/2023

<u>Background</u>

Developing techniques during the first two years of my PhD research afforded new approaches to analysing the data collected in 2021/2022. The data collected consisted of individual in-depth interviews, online survey (questionnaires), and Twitter data. The interview and survey data were initially collected to be analysed using NVivo software. During the progress of the research, Artificial Intelligence methods were explored to fully integrate a Grounded Theory approach for an initial unbiased analysis of the datasets.

<u>Method</u>

Natural Language Processing/Topic Modelling (Jones 2021; Jones, Doane and Attbom 2021) provides a cutting-edge alternative for initial exploration, categorisation (clustering) and topic labelling. This method uses Latent Dirichlet Allocation (LDA), a statistical analysis of the data. During the analysis, temporary models are created based on the statistical analysis, which pre-orders and pre-labels the data. The models are not created for publication or as a form of Machine Learning (as a basis for an AI analysis in future). Setting a “seed” during the process guarantees the replicability of the results. But at this stage, no modelling is used to provide the tool itself. In this regard, the results are directly observed by the researcher and further analysed manually. A statistical summary and categorisation of the Topic Modelling process are similar to the coding resulting from using NVivo.

Anonymised quotes from participants will be used during the publication of the summarised results. The participants will not be identifiable as the individual story, and their location of the favourite place will not be possible to associate with individual survey participants. Interview participants have given consent to being named and not anonymised as the in-depth case studies reflect the personal life story of the interviewees.

In no case will the model itself be shared with a third party. In the case of a paper submitted to a journal using Topic Modelling (LDA) for the initial analysis of survey data, the modelling was purely statistical to provide an unbiased insight into the data. The result was later manually coded and categorised.

Potential Method Development

As mentioned above, the LDA statistical model analysis is not currently used in a Machine Learning approach to create supervised models that autonomously categorise a new dataset of survey responses by a third party. In the case that a supervised model is created based on the provided survey or interview data (which means the data are used to train a Machine Learning model that is capable of categorising new data), this model will not be shared with a third party to use the model based on real-person information to create policies or regulations. The methodology and outcome will be published as proof of concept. The original model will not be shared, which will acknowledge that no consent has been sought to use the data for the actual real-world decision-making process.

Conclusion

Ethical concerns associated using of Artificial Intelligence methods can arise when models are based on the information of living people and the models are used to influence or create policies, regulations or agendas actively. In this case, the participants would need to consent to such use. However, the methodology of my research consists of two stages:

a) Unsupervised Topic Modelling using the statistical DLA method that creates temporary models for each analysis. The models are not used for any other purpose than creating topic clusters to analyse the respective dataset.

b) Supervised Topic Modelling to create a model trained on the datasets and used to assess the model's performance in assessing new data. In this case, the proof of concept will be shared but not the model itself.

References

Jones, T., Doane, W. and Attbom, M. (2021). *textmineR: Functions for Text Mining and Topic Modeling*. [Online]. Available at: https://CRAN.R-project.org/package=textmineR [Accessed 19 August 2022].

Jones, T. W. (2021). Topic modeling. [Online]. Available at: https://cran.r-project.org/web/packages/textmineR/vignettes/c_topic_modeling.html [Accessed 31 July 2022].