What makes for effective and meaningful online parliamentary public engagement?

Evaluating the UK Parliament’s digital engagement activities

Nicole Dezrene Rhea Nisbett

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
School of Politics and International Studies

November 2020

Acknowledgements

There are many people without whom this research would not have been possible. The White Rose Doctoral Training Partnership for awarding me a scholarship to undertake this PhD. The Digital Engagement team, specifically Laura Bristow, Matt Ringer, Laura Wickens, and Ben Pearson, for collaborating with me over the last three years and educating me on the work they do to interact with the public. The SoBigData Research Infrastructure for funding my research visits to the University of Pisa, and Giulio Rossetti for his guidance. I'd especially like to thank my supervisors, Dr Viktoria Spaiser and Professor Cristina Leston-Bandeira for their steady support, motivation, and trust in me completing this research. Their encouragement when my motivation was waning and their understanding of my many squirrel rants were much-needed life rafts.

I'd also like to thank my family and friends. The 9th floor of the University of Leeds Social Sciences building and the people who made this experience so enjoyable; Ruth, Naba, and Ramya for the 3-hour lunches and nonsensical conversations; Ahmed and Daniel for being my surrogate big brothers, travel buddies, and fellow outcasts in the department; and previous members of the WMLC Ben and Alex for their occasionally helpful advice. A special thanks to Professor Dorothy Monekosso for her invaluable mentorship and introducing me to a new field.

And finally, to my parents for their many sacrifices and unwavering support, this was always for you.

Abstract

This research primarily aims to develop evaluation methods to effectively harness citizen input from large unstructured data generated automatically through digital engagement activities. It is an interdisciplinary and collaborative project with the House of Commons which combines social science and data science to analyse the online engagement activities of the UK Parliament. Digital Engagement teams within Parliament have introduced various ways of engaging with the public online including consultations and digital debates. These have been popular since they started in 2015 but attract too many responses for staff to process manually and to get a clear picture of what the public is saying. I use machine learning and text mining approaches to analyse the data gathered by Parliament to summarise and reveal the network of participant interactions so Parliament can have a more informed idea of who is participating within which social/ideological clusters. This shows a public who have a diverse set of views but can be influenced based on the channel and type of engagement they are participating in.

As the Members of Parliament are crucial to the engagement process, any way to encourage and facilitate their use of the online engagement is vital. Without input from officials overtly showing that they have listened to and incorporated the public's opinions into their decisions, the online public engagement efforts from Parliament could be seen as insincere to many of the public. With this in mind, another aim is to explore how public opinion derived through the online engagement activities can be meaningfully incorporated into policy making. This entails working with different teams in Parliament to understand exactly how policy-makers are currently using the outputs of online engagement and how this can be improved. I conduct demonstration tests to test the methods of evaluation developed during the research and find that while these can be applied to digital engagement activities successfully to gain insights from the public, responsibility remains with the institution to ensure internal processes are equipped to make use of the public's views.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Meaning
BSL	British Sign Language
DTM	Document-Term Matrix
EAC	Environmental Audit Committee
EFRA	Environment, Food, and Rural Affairs
IDF	Inverse Document Frequency
IPU	Inter-Parliamentary Union
LDA	Latent Dirichlet Allocation
NRC	National Research Council Canada
POS	Part-of-speech
RT	Retweet
TF	Term Frequency
WPU	Web and Publications Unit

Chapter 1 Introduction

1.1 Summary of research

Discussions of contemporary politics have focussed recently on the so-called "crisis of disengagement" bemoaning the increasing disconnection between the public and holders of political authority (Norris, 2011; Webb, 2013; Flinders, 2015). On the other hand, political debate takes place increasingly on the internet and often within disconnected and polarised "bubbles" (Brundidge, 2010), which escalate disagreements and do little to promote constructive discussion and compromise. Both phenomena can challenge established democratic systems and contest traditional representative democracy practices, if the public does not feel sufficiently represented by policy makers and if democratic compromises become harder to reach. In response to these transformations, policy makers have sought new ways to engage with citizens, increasingly making use of digital technologies. Digital debates involving MPs and citizens prior to parliamentary debates are one example of such attempts developed recently by the UK Parliament.

However, current analysis of these activities is limited to descriptive statistics which measure the quantity rather than quality of engagement. This research is collaborative with the House of Commons and will address the evaluation of effectiveness and impact of Parliament's online public engagement activities. The House of Commons Digital Engagement team, which sits under the umbrella of the Participation Team in the UK Parliament, have collaborated on this project to develop a better understanding of the data created through their online engagement activities. A thorough computational analysis of the people interacting online with Parliament and their input will be completed to understand their views and feelings towards specific topics, as well as the underlying digital community clusters formed as a result of their participation. The development of demonstration tests to investigate the effect of moderation as well as the use of alternative participation platforms will help to fine-tune understanding of digital participation. This will delve deeper into the evaluation of digital engagement and identify the importance of factors such as specificity of platforms and of the processes behind digital engagement initiatives. To aid in the evaluation of digital engagement, I create a web-application TheGist which summarises and subsets online discussions based on the main themes discussed and sentiments expressed by online participants.

I find that while it is important to address the external barriers to engagement such as ease of public access to engagement tools, the internal barriers faced by parliamentary staff such as lack of resources and procedural processes are vital to have a successful and sustainable engagement strategy. Concrete understanding of the context in which the engagement is happening must be considered before diving into the overwhelming choice of digital tools. This PhD's contribution is therefore two-fold: an analysis of parliamentary public engagement to evaluate citizen input in the context of digital engagement with the UK Parliament, as well as the development and introduction of alternative digital tools for practical implementation of insightful digital engagement sessions. This includes a bespoke web application, TheGist, that allows officials to effectively harness citizen input.

1.1.1 Identifying the research problem

Key authors in the field of public engagement introduce various definitions of engagement. They represent the dimensions of engagement as different levels each with their own objectives in terms of their communication with the public. Some prioritize providing information to citizens about the institution, while others seek to consult the public on specific issues. Models such as those from Lenihan (2008); OECD (2009); Rowe and Frewer (2005) help to conceptualise public engagement depending on the desired end-result, i.e. to consult the public or to generate discussion. The question of whether the institution's motivations for engagement are genuine to encourage the public to get involved, or simply to give the appearance of engaging with citizens is also raised in Arnstein (1969). However, for the purposes of this research a new conceptualisation is proposed which treats engagement as a spectrum, differentiating whether or not there is a need for input from the public in the engagement activities of an institution – specifically the UK Parliament. This new spectrum is developed through analysis of the current literature, incorporating specific elements of existing conceptualisations.

Through its previous public engagement strategies, it becomes apparent that Parliament also treats public engagement as a combination of different dimensions. For example, their first public engagement strategy from 2006 to 2011 had three aims: to inform, to educate, to promote (Walker, 2012). This helped Parliament to embrace digital technologies such as the internet to improve their website, as well as an increase in media broadcasting to inform the public of business. An education centre would also be introduced later which would be focussed on educating the public on various aspects of Parliament such as the history of parliament and parliamentary processes, as well as encouraging visits from schools and providing resources for teachers of parliamentary studies. However, at that time the focus was heavily on one dimension of engagement, namely informing, and neglected the participatory aspects.

Nevertheless, while not fully engaging the public in policy, Parliament's focus on just the one-directional information flow out of the institution did provide the foundation for their second public engagement strategy of 2011-2016. This placed much greater emphasis on the participatory dimension of engagement which sought to involve the public in a way that went beyond consuming information from the institution. Select Committees within Parliament also gave prominence to public engagement by introducing it as a 'core task', just as had been done with scrutiny of government in 2010 (Liaison Committee, 2015). However, these developments do not necessarily lead to more positive attitudes amongst the public with respect to the Parliament. Although 73% of the public claim to value Parliament, only 30% are satisfied with the way it works (Hansard Society, 2017, p. 7). Public engagement has not always been a key priority within the UK Parliament, and even at the education level, one of the most recent Hansard Society Audits of Political Engagement reports that only 49% of people claim to know "a fair amount" about Parliament (Hansard Society, 2018, p.36). That being said, a change is most definitely underway within the UK Parliament which places this project in a unique position, allowing it to be a part of the change as it unfolds.

Further evidence of Parliament's commitment to digital engagement is through the Digital Democracy Commission. This was created and chaired by Speaker Bercow in 2014 and made up of experts from a range of sectors including academia, charity, public and private sector (Digital Democracy Commission, 2015; Good Things Foundation, 2017). The commission released a report outlining the ways Parliament could make use of digital tools to engage with the public. Leading by example, a range of online channels were used to engage with the public in this report including social media, web forums and surveys, and live-streaming of meetings (Digital Democracy Commission, 2015, p.77). This report also put forward a set of recommendations, several of which have been taken on board and are now used in the day-to-day running of Parliament. One of these recommendations is digital debates, several of which are analysed in Chapter 6. Digital debates are proposed by Members of Parliament (MPs) and have been running since July 2015 (Parliament.UK, 2017).

One area highlighted within the Digital Democracy Commission (DDC) report was the difficulty in using online forums for public consultations due to the large amount of data this would create. The number of resources needed to successfully manage a sustained online forum is vast in terms of staff, time, and expertise (Digital Democracy Commission, 2015, p.49). The DDC report is cautious not to suggest or recommend a subject-wide public forum which could generate tens of thousands of messages for any area of government, further highlighting an internal barrier to engagement. Although forums such as this exist in other industries and countries, the need and expectations from the users for the moderator to read each message and offer a thoughtful response is not as important.

Scalability is another element of this research, and the consequence of not factoring it in can add to the feelings of distrust and disengagement on the part of the public who may feel as though their opinions are not taken into account. This worry is also raised in the World e-Parliament report (Inter-Parliamentary Union, 2016) which found that inadequate staff capacity (49%) and lack of ICT knowledge (43%) were big challenges to parliaments around the world with regards to using ICT. This highlights an internal barrier to public engagement specific to Parliament – people do not only wish to voice their opinions; the majority want to be acknowledged and see a change as a result. The problem that arises with this and leading to the DDC's understandable caution is the time it would take to read each comment in a forum.

Likewise, one must determine a way for parliamentary officials to make the best use out of the comments given by the public to contribute to decision making. Consequently, by not having the correct infrastructure in place (be that digital or human) the online engagement initiatives may do more harm than good. Citizens may feel that their opinions are not being valued or taken into consideration by parliamentary officials and that they are wasting their time in participating with the Parliament. As a result, they may develop a negative opinion regarding the trustworthiness of the institution and be reluctant to participate in the future. Smith (2009) also recognises these consequences of a lack of engagement by officials and what genuine influence discussions could have over decision making. Therefore, addressing the cost-benefit analysis of effort (of using online tools) versus returns (in terms of contribution to legislation) for the Members as well as the public is absolutely crucial to understanding the value and effectiveness of online digital engagement initiatives. Previous studies have found that when not adequately incorporated onto the parliamentary process, the actual contribution the participation sessions have on legislation is slim (Leston-Bandeira and Thompson, 2017). This project will aim to begin to remedy this barrier to engagement using text analytic tools and techniques[1]. Throughout this thesis, textual data is collected and analysed automatically, culminating in the development of a purpose-built web application which allows officials to complete this analysis themselves without the need for coding. The development of this application addresses therefore a major challenge for parliaments and provides them with a tool to more effectively understand and integrate inputs generated by digital engagement into policy-making. The tool also serves as a major contribution of this thesis.

The impact of digital tools has been powerful and, in some ways, sudden over the past few decades. Parliament, like most institutions produces a vast amount of data, be that Hansard reports, television broadcasting, or their social media account data. Summary statistics and descriptive analyses can be derived from this data to provide an overview, but these do not fully explain what is going on. Innovations in text analytics (Adeva *et al.*, 2014; Aggarwal and Zhai, 2012), topic modelling (Blei, 2017; Blei, Ng and Jordan, 2003), and network analysis (Bedi and Sharma, 2016; Dean, 2018; Smith *et al.*, 2014) allow us to really appreciate the stories behind the data and understand its value in a wider context. As such, another input of this thesis is the development of a bespoke analysis tool, TheGist (Chapter 7), specific to the data sources created by Parliament, allowing for a clearer understanding of the contributions of the citizens who participate online.

[1] This will be explained further in Chapter 6

Of course, when any discussion is taken to the online medium it can introduce unique problems. The notion of technopopulism (Coleman and Gotze, 2001, p.8) where the loudest voices dominate the discussion, and extended self (Belk, 2016) where users portray themselves differently online is something to consider, especially when the option of anonymity is made available. This can introduce the problem of so-called 'trolls' or even instances of cyber-bullying, something which people would not do in face-to-face conversations. Moreover, while people do indeed recognize that social media can help with bringing ordinarily underrepresented groups into the political, there is also an awareness that conversations can become more divisive and even superficial than before (Hansard Society, 2018, p.15).

The question of an echo chamber versus a public sphere has been raised in Colleoni, Rozza and Arvidsson (2014) to differentiate different types of online communication. An echo chamber relates to a conversation in which opinions on a certain topic are shared by participants and are often one-sided leading to an echo effect of views with little to no debate. This is linked to homophily where the clusters of participants have a higher likelihood of communicating with each other than with participants outside of their cluster (Asher, Leston-Bandeira and Spaiser, 2017; Zalmout and Ghanem, 2013). On the other hand, a public sphere scenario has the participants exposed to differing opinions to be debated. The latter scenario is something to aspire to as most people have a tendency to communicate with or seek out people more similar to themselves (Papacharissi, 2002, p.23; Carpini, Cook and Jacobs, 2004, p.335). However, some argue that the presence of homophily and echo chamber discussions provides a better environment for true deliberative democracy (Mutz and Wojcieszak, 2009). The existence of either of these scenarios within the online discourse of the UK parliamentary debate can shine light on the behaviour of politically interested citizens. Something else to consider is the representativeness of the data source is something to consider when drawing any conclusions. Only 19% of people claim to have visited a political social media account, however this figure increases to 29% for 18-34 year olds (the largest demographic online) (Hansard Society, 2018, p.29).

Controlling for the potential problems above, the use of digital debates as a means to encourage participation could be a difficult way to improve the public's perception of the value in participating with the institution as only "32% of the public think that debating important issues in the House of Commons is an important way for MPs to spend their time." (Hansard Audit 14, p.7). Nevertheless, even with these shortcomings social media and other online platforms remain a valuable channel of communication whose use is becoming increasingly widespread. This is also becoming noticed in Parliament with Robert Halfon MP (also on the DDC) declaring that more must be done in the online sphere regarding public engagement (Halfon, 2018).

Another important consideration to bear in mind when analysing online debates is the concept of digital divide (specifically in relation to online political participation), discussed by various researchers. Specifically, Epstein, Newhart and Vernon (2014) argue that there exists two levels of divide, the first relating to accessibility to the internet and the second to motivation. Another conceptualisation from van Dijk (2006) categorises the divide by the skills needed; namely operational – the use of the core hardware and software, information – the ability to search for information online, and strategic – the ability to use IT to achieve goals. As of February 2020, 96% of households in the United Kingdom have access to the internet (Office for National Statistics, 2020). This proportion is much higher than the global average of 59% (Statista, 2020) suggesting the digital divide is not as prevalent in the UK in comparison to other countries. This means that only a small sector of society is restricted from engaging with the Parliament online as a result of lack of access alone. This does not mean that there are not other limitations which prevent the public from participating, for example, the second level of lack of motivation as described by Epstein, Newhart and Vernon (2014). This could also be a result of a perceived lack of representativeness by MPs to their constituents. Although this project will not be focussed on addressing or closing the digital divide, this will still be an important issue that the study will consider and ensure any conclusions have taken the discrepancy between the 'haves' and the 'have-nots' into account.

Therefore, while studies have been done regarding the components of public engagement, we still lack an understanding of the level to which digital tools can improve or worsen the effectiveness of engagement. There are many factors to consider when using online methods including representativeness of the online audience, the amplification of certain personality traits, and the resources needed to effectively take part in an online participation session, to name a few. Combined with a parliamentary institution which is trying to introduce changes to become more public engagement friendly, this raises the need to devise a strategy in which the use of digital tools can enhance public engagement and be of benefit to MPs and citizens alike; and, in order to achieve these, the need to be able to better understand digital engagement data and be able to evaluate it within parliamentary timings and structures. Furthermore, the different dimensions of engagement should be taken into account when analysing any online engagement strategy.

1.1.2 Research Questions and Methodology

Due to the nature of this project, the methodological aspects form a core contribution of this thesis and will be fully detailed in a substantial Methodological Framework chapter within the thesis (Chapter 3). Taking the research problem into account, the three research questions are:

1) How can we define and evaluate the effectiveness of UK Parliament's approaches to online public engagement?
2) What is the nature of citizen input in digital discussions initiated by the UK Parliament?
    a) What can we learn about people's views on the issues raised in these digital discussions?
    b) What can we learn about the participants involved in these digital discussions?
    c) What can we learn from participants' interactions during these digital discussions?
3) How can the citizen input be utilised in a meaningful way to inform policy making?

The first question requires a definition of 'effective' – namely understanding what Parliament aims to achieve with their engagement activities. A continuing argument in this thesis is the importance of clarity and managing expectations with regards to how engagement activities are conducted, and how they are evaluated as a result. It is important for the team of officials implementing an engagement activity to set goals and priorities in terms of the type of engagement (i.e. informing vs. consulting) they are conducting. It is also important for the public participating to be made aware and comprehend how, if at all, their contributions will be used. Through understanding what the priorities are for an engagement activity, I can begin to develop methods of evaluation, in turn determining how effective the activity has been relative to the original goals. This question will address and identify the different elements of engagement, both the activities, aiming to simply disseminate information such as the use of Twitter, and the activities which aim to gather opinions from digital debates. Through working not only with the Digital Engagement team, but also other teams with a remit for online public engagement in Parliament I can begin to understand how the definition of effectiveness varies depending on who is leading the activity. This definition of effectiveness will be crucial in future analysis and will inform the development of demonstration tests.

Once I have discussed how engagement can be effective and established methods for evaluating engagement activities, I can focus on the second research question which focusses on understanding what the participants contribute to digital discussions in terms of the issues raised, their demographic attributes, and how they interact with each other. There are gaps in our understanding of how useful natural language processing can be for digital engagement teams in parliaments, with the goal to effectively harness citizen input. Therefore, this question explores the computational techniques which can be applied to public engagement data, specifically comments from social media platforms such as Facebook and Twitter. Measuring both what users are saying using text mining methods such as topic modelling algorithms (Hong and Davison, 2010; Wang and Blei, 2011) and sentiment analysis (Nielsen, 2011b; Pu, 2017; Sen, Rudra and Ghosh, 2015; Verma *et al.*, 2011), and how those users are interacting with each other using social network analysis and community detection algorithms (Bedi and Sharma, 2016; Papadopoulos *et al.*, 2012) provides a comprehensive account of a discussion. Another aim is to identify the different backgrounds of these users (be that socio-economic, political leaning, or educational) as it is important to determine whether Parliament has been successful in reaching a varied audience (Sloan *et al.*, 2015). The inclusion of location data for geo-spatial analysis would also provide MPs and officials with estimated constituent-specific issues that they may not be otherwise aware of. In the datasets I have in this project, constituent-level geographical data is not available, but I can still gain city-level information from the participants on social media. Through this layered approach, I can create a picture of who the participants are and what they are saying. Crucially, this can be achieved for online discussions with thousands of comments which would usually be analysed manually, taking up many resources from parliamentary staff. Introducing computational methods also allows Parliament to conduct more digital engagement activities by helping to break down some of the internal barriers to engagement, such as lack of time for analysis.

The third question concentrates on understanding what effect these public participation initiatives are currently having on policy making, as well as how this can be enhanced. One would expect a public engagement activity to have a clear understanding of how the views of the public will be integrated into parliamentary business, which will also facilitate closing the feedback loop with the public. By definition, this is more important for those activities which seek to gather views from the public rather than just provide information. While also having different priorities and measures of success, various teams undertaking public engagement may also have different ways of implementing the findings from engagement activities. Therefore, in Chapter 4 I examine the organisational structure of teams with a remit for online public engagement in the UK Parliament to have a better understanding of how each team deals with the activities. Currently the UK Parliament uses social media platforms to conduct their engagement activities. While these have the advantage of having an in-built community of users, they were not designed with online discussions in mind. I therefore collaborate with select committees to experiment with an alternative platform, Discourse, specifically created to encourage digital discussions. In comparing several platforms, I can understand how Parliament is using the platforms for different types of engagement and measure how the public interacts on different platforms. These insights can help teams understand how the use of one platform can influence the types of responses and discussions one might receive during an engagement activity, and help inform future engagement sessions. The collaboration with select committees during the Discourse demonstration tests also provides insight into how a different type of citizen input can be used in committee inquiry reports and linked with other parliamentary activities and data.

All three questions require the use of programming and analysis software to draw out the important insights from the data, mainly using R. The data in question will come from Parliament's current public engagement activities, primarily on Twitter and Facebook, but also external platforms used by the Digital Engagement Team in the House of Commons (described in section 3.1). The insights derived from this data could prove invaluable to the UK Parliament in relation to the three questions above, not only helping them to understand their data and reach of activities, but also encouraging policy-makers to embrace online public engagement in its entirety. Bringing these computational methods into the day-to-day running of Parliament will also help the institution adapt to new technologies allowing it to adequately manage and welcome the increasing demands introduced by digital engagement rather than seeing them as a barrier. Explaining these methods in such a way that non-programmers can easily use and interpret the results will also be a valuable contribution of this research and ensure the sustainability of the techniques introduced. The insights gained from this collaborative project can also be applied to the existing literature on parliamentary engagement and how institutions can learn from and make the most use of their interactions with the public. On the other hand, this thesis also contributes towards identifying internal parliamentary barriers which may be impeding the process of digital engagement.

1.2 Contributions, limitations, and impact of research

The collaborative nature of this PhD places the research in a very fortunate position. As the pursuit of research questions and the ongoing progress involved the House of Commons, including access to the Parliamentary Estate for the duration of the study, the project was able to adapt to changes in processes as they happened. This ensures not only the pursuit of scholarly research questions, but also the development of research outputs that are aligned with the practical challenges faced by Parliament, being therefore as impactful as possible. Nevertheless, maintaining the rigour and independence of the research was paramount and I was careful to ensure the collaboration did not hinder research independence.

The impact of this research has several facets. Firstly, by understanding exactly who the current engagement practices in Parliament are targeting and what types of conversations and communities they attract, the Digital Engagement team especially are able to target their future digital debates and discussions accordingly. Having interacted with several other teams including Parliament Digital Service, Petitions Committee and Select Committee Engagement teams, the outputs of this research can reach a wider audience within Parliament. Large data sources have been shared with me, which facilitated a thorough analysis of the users who interacted with all the Parliament-owned and associated social media accounts – information which is crucial for informing future initiatives.

Secondly, introducing a digital tool[2] to analyse data derived from public engagement at scale, allows Parliament to control incoming data in a way that is currently not possible without a significant amount of staff hours. By facilitating the analysis of large-scale online engagement activities, teams can embrace all aspects of engagement without fear of not meeting expectations due to inability to process large volumes of public input. This tool allows Parliament to remain more up-to-date with innovations in data science, specifically natural language processing, to get more insights out of their data. This therefore contributes to the productivity of the teams, allowing them the time to focus on other tasks, as well as encouraging some Select Committees to embrace digital tools in their inquiries and when gathering evidence. As previously mentioned, they have tended to shy away from using certain forms of social media due to the fear of a greater number of comments than they can handle. Therefore, by digitising the analysis process for large volumes of text, they can be more open to new forms of engagement.

[2] https://github.com/NicoleDNisbett/TheGist

Finally, there currently exists no public Parliament-wide organisational structure of teams dealing with online public engagement. Through speaking with various teams, it becomes clear that the lack of cohesion and conversation regarding the types of data held by each team causes a situation in which one team's processes are held back by being unaware of another team's data. Creating an organogram that outlines these different teams and how they relate with each other, clarifies exactly who is under the remit of online public engagement, and therefore who may possess data needed by others. This is explained further in Chapter 4.

So far, this section has highlighted this project's impacts on the UK Parliament, however there is also impact on the contribution of knowledge in this area of parliamentary engagement. Through analysis of literature and practice of engagement, I introduce an alternative UK Parliament-specific spectrum of public engagement. This addresses engagement from an angle of the importance of the public's input as well as the direction flow and source of information. It places particular attention on differentiating public engagement based on the priorities of the institution carrying out the activity. On one hand, providing information to the public so they are informed and educated about the business of Parliament is often conceptualised as the first stage or step of public engagement as described by Arnstein (1969); Leston-Bandeira (2012); Kalampokis, Tambouris and Tarabanis (2008). The other side of engagement prioritises seeking views from the public and integrating them in some way into the political and decision-making processes. However, where many scholars treat these dimensions of engagement as distinct from each other, I argue that they are equally as important in the engagement process. Furthermore, many engagement activities can include more than one dimension, for example, information and consultation with both being vital components to the success of the activity. Therefore, the spectrum of engagement (introduced in section 2.2) conceptualises the different dimensions of engagement as equal in importance while also maintaining and recognising their individual contributions to the public's understanding and engagement with Parliament.

Another contribution of knowledge of this project lies in the evaluation of public engagement activities. Section 2.5 highlights how the difficulties of evaluating engagement are caused by its many dimensions and lack of a systemised conceptualisation across practitioners of public engagement. As a result, many different evaluation measures have been developed depending on the aims of the institution carrying out the activities. I introduce various methods to evaluate engagement activities based on the reasons for conducting the engagement session as defined by the different dimensions (i.e. to inform or to consult). These evaluation measures are focussed on employing techniques in text mining and social network analysis to understand how well the activity has performed according to the initial priorities of the institution. Examining engagement in this way firstly ensures that the institution is clear from the outset about exactly what it intends to achieve from conducting an engagement activity, secondly ensures that they have a meaningful way of evaluating the activity, and thirdly ensures that the public's expectations of the activity are managed and they do not become un-encouraged from engaging in the future due to lack of clarity. Ultimately, this research adds to the literature of parliamentary public engagement, specifically online.

While the contributions of this research are strong, there are also limitations found in conducting interdisciplinary and collaborative research. The UK Parliament is a very risk averse institution, which is cautious when introducing any new processes or technology. Furthermore, I was often dependent on Parliament regarding the type of data I was able to analyse and they mediated the access to this data. For this reason, some analysis of data was not carried out. Data such as names of users on Facebook which could be used to estimate the gender distribution of participants on digital discussion cards is not analysed as Parliament felt it was too intrusive to the participants[3]. Therefore, only the data which would be used by them was analysed to ensure there was a need and purpose to the work.

There are some methodological limitations of the analyses presented here that should be explicitly stated. Some of the analysis could not be validated, mainly due to a lack of data that would allow for validation. The terms of the ethical approval for this work and the collaborative nature, with UK Parliament being particularly concerned with user privacy, meant that certain types of data, in particular user data, was not available or could not be extracted. Furthermore, the interoperability of the results was of particular importance in the context of this collaborative project, which meant that sometimes interpretability played a greater role in choosing for instance topic models, than metrics, which however limits the validation of the topic model results. Ultimately, this is not a methodological project, where the goal was to develop new approaches for analysis, rather the goal was to make use of well-established methods and apply the in the context of this collaborative work in a way that makes the result accessible to users without a technical background. I was also somewhat limited by the collaborative nature of this project which determined the type of data I was able to extract. The UK Parliament is a very risk-averse institution with major concerns regarding privacy, data collection, and data security which I needed to adhere to. This is explained further in section 4.3.

There were delays in getting this application used by the teams it was created for, due to concerns over data security and the technical details behind the use of the application (explained further in section 4.3). Furthermore, the demonstration tests conducted in section 6.5 lasted for only one week on average. I would have ideally preferred these to be longer, however I had to adhere to the schedules of the parliamentary committees involved. Longer discussions would have allowed more time for participants to be involved in the platforms and perhaps alter the types of conversations observed. The limitations with respect to data and methodology will be discussed in the methods section (Chapter 3).

Nevertheless, the collaboration with an external partner and the emphasis on methodological techniques surrounding the evaluation of engagement puts this project in a very unique position. I was able to immerse myself into the teams at the forefront of parliamentary engagement activities in the UK and understand how their internal processes impact how they conduct engagement. Through this I was able to introduce methodological techniques to the evaluation of engagement activities and create a truly collaborative and interdisciplinary research project.

[3] Personal communication, Westminster, November 2018

Chapter 2 The multiple dimensions of public engagement

This chapter explains key concepts and perspectives regarding what public engagement is and how it has developed. It aims to identify a framework to evaluate digital public engagement activities, specifically in the case of the UK Parliament. As is shown below, a single definition of public engagement is hard to find due to the many actors involved with these types of activities, so I explore the different dimensions of engagement using real-world examples to evidence the theory. Before I get to analyse parliamentary public engagement, I need to discuss core ideas and scholarly contributions on this concept. I begin in section 2.1 by exploring why public engagement is necessary and why it is conducted in different ways. In section 2.2 I introduce a new interpretation of digital engagement which incorporates different online engagement activities categorised into various dimensions. This spectrum of engagement is outlined in this chapter and builds on the existing literature on public engagement. These dimensions are distinguished by those which require input from the public and those which do not. This distinction is developed through examination of several existing models of engagement with each focusing on a particular aspect of public engagement. This interpretation of engagement as a spectrum includes activities which are specific to those conducted by the UK Parliament but can also be generalised to other institutions carrying out digital engagement.

Sections 2.3 and 2.4 explain the dimensions of engagement in the spectrum in more detail, including examples of how this interpretation of engagement has been practised in the UK Parliament. One key aim of this thesis is to understand how public engagement is evaluated and explore different ways to improve this evaluation for officials who manage digital engagement activities daily. To do this, section 2.5 explores some existing methods of public engagement evaluation used in various contexts around the world to see how engagement is currently measured and ways for improvement. Finally, section 2.6 explores examples of select committees using public engagement in different settings to contribute to their inquiries and work in scrutinising the government.

2.1 Why public engagement?

Understanding any issues regarding public engagement first requires an understanding of why engagement is important. The value of this type of research is shown through various studies showing there has been a notion of distrust of Parliament or even hatred of politicians by the public (Hay, 2007; Dalton, 2004; Norris, 2011). Other researchers posit that the public fall into different categories based on their reasons for disengagement (Webb, 2013). Some wish to be more involved with democratic processes and feel that the existing system does not suit their needs (dissatisfied democrats) while others are happy to simply leave the decisions up to the elected (stealth democrats) (Webb, 2013). Most research so far has focused mainly on the public's perspective, we still lack research on how institutions deal with and consider engagement form the public. For as important as it is to understand why and how the public engages with political institutions, it is also important to consider the perspective from the institution.

One motivating factor for parliaments wanting to engage with citizens are votes. The UK Parliament is a representative democratic institution which relies on the votes of the public to elect its members (Dorey and Purvis, 2018). Every 5 years or so, the public vote in their millions for who they believe should form a government and pass laws. Voter turnout is therefore a common measure of how much the public is involved with politics, and as a result Parliament. In the UK, election turnout has been steadily rising since its 2001 low point of 59%

and the Institute for Democracy and Electoral Assistance report that more people voted in the 2017 general election with 69% compared to 66% in 2015 (IDEA, 2018). And although this was slightly down in 2019 this was still very high, particularly bearing in mind the idiosyncratic nature of this election, when the UK people had been called to vote three times within just four years. Voter turnout is also important with regards to representation of the electorate. If only a small percentage of the population turn out to vote, the claim of representing the views of the people is weakened. Furthermore, there always exists a proportion of the population who consistently do not vote. Those are often the marginalised and underrepresented sections of society (Goldfinch, Gauld and Herbison, 2009; Norris, 2001), and so reaching these people is an important task of parliaments around the world (Dalton, 2004; Norris, 2011) .

On the other hand, trust in the UK Parliament had taken a steep drop in the wake of the 2009 MP expenses scandal, as well as the Government's U-turn on the University fees and the NHS in 2010 (Lee and Young, 2013), with the percentage of people claiming to trust Parliament falling from 37% to 18% between 2007 and 2009. However, as Figure 1 shows this value was steadily rising once more and stood at 34% in 2017 but has since dropped again to 21% in 2019. The data also shows how trust in Parliament and Government follow similar trajectories. Any rises or falls in tendency to trust Government are mirrored by Parliament generally between 2 and 4 percentage points higher (or lower in the case of 2009) than Government. This reinforces the closeness in perception of the two institutions by the public and how this may make Parliament's job of differentiating themselves from Government more difficult. And as I will demonstrate in subsequent chapters, this is a particularly important point to bear in mind when it comes to implementing public engagement initiatives with Parliament and evaluating these.

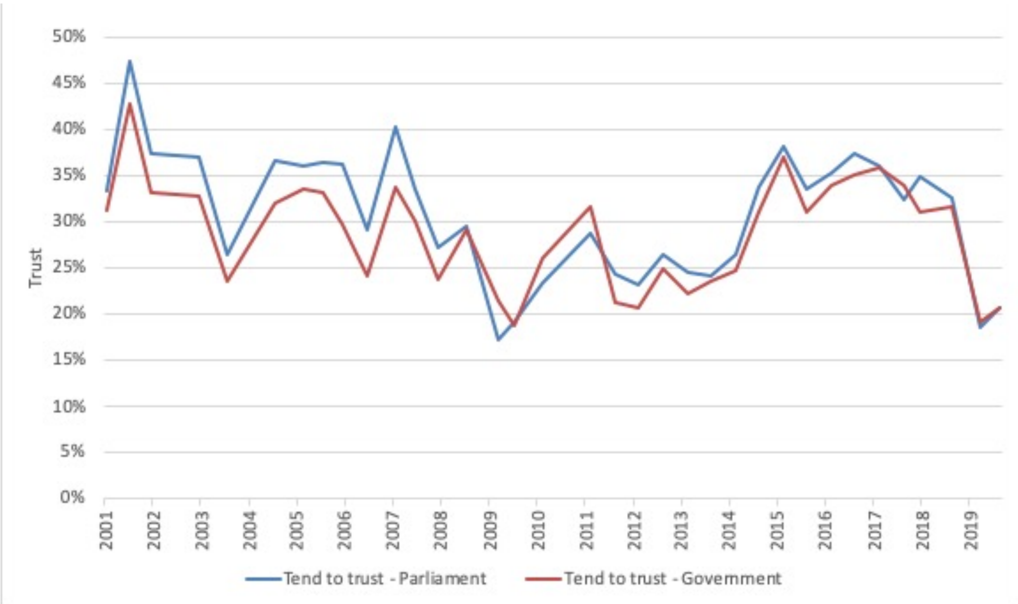


Figure 1: Trust in Parliament vs. Government 2001-2019 (Eurobarometer 2019)

But what happens in-between the general elections, and why? Studies have shown that the public wish more and more to be involved in politics more often than every 5 years, at election time (Dalton, 2004; Norris, 2011). Therefore, creating clear public engagement strategies which allow for continued involvement with the business of Parliament should be the focus of institutions. This chapter explores how engagement can be broken down to provide a starting point for these strategies.

2.2 Not all engagement is created equal

This subsection reviews how public engagement is understood by different authors and parliamentary actors, the different aspects which incorporate public engagement, and introduces a new interpretation which marries the conceptualisations of several authors in the field. This is an important stage which adds to our understanding of the different components involved in public engagement, the difficulty in defining engagement, and how they can in turn influence the evaluation of engagement activities.

The definition of public engagement can differ depending on the many factors and angles involved, and several authors have raised the issue of multiple facets to be considered in categorizing public engagement. As Leston-Bandeira (2012) notes, "Public engagement covers a very wide range of outlets and activities with different purposes from information to participation in public policy." (p.418). In the Modernisation Select Committee's *Connecting Parliament to the Public* report in 2004, the recommendations to Parliament to advance public engagement involved everything from educational initiatives and visitor activities, using the website and digital broadcasting to opening up to the public more, and the increased use of petitions as they "represent a potentially significant avenue for communication between the public and Parliament." (Modernisation Committee, 2004, p.3-8). Several areas of interest are also raised in Walker (2012) when outlining the UK Parliament's first 5-year strategy developed as a response to the Modernisation Committee's recommendations. This strategy lasted between 2006 and 2011 and outlined three main aims: "to inform the public about the work and role of Parliament; to promote Parliament as an institution and explain why it should be valued; and to listen to the public by seeking and responding to feedback." (p. 272). This helps towards beginning to categorise the different aspects of public engagement, in the latter's case to inform, promote, and listen, while Leston-Bandeira (2012) provides a 5 step process ranging from simple provision of information at one end to full citizen participation and intervention in parliamentary business at the other (p. 418).

The lack of coherence regarding a versatile public engagement strategy in various parliaments around the world provides further evidence to support the difficulty in defining engagement. In their, '*Parliaments and Public Engagement*' report, the Hansard Society (2011) recognises the problem of evaluating such strategies lies in the fact that parliaments' strategies are rather sparse and unstructured to begin with.[4] The parliaments of Australia, Denmark, the UK, and Wales are highlighted as having the most comprehensive public engagement strategies. A key underlying theme of these strategies is the clear focus or 'target market' of their populations. For example, Denmark's population is split into professionals, communicators, and citizens, while Wales' categories are those who must know, those who need to know, and those who would like to know (Hansard Society, 2011, p.77).

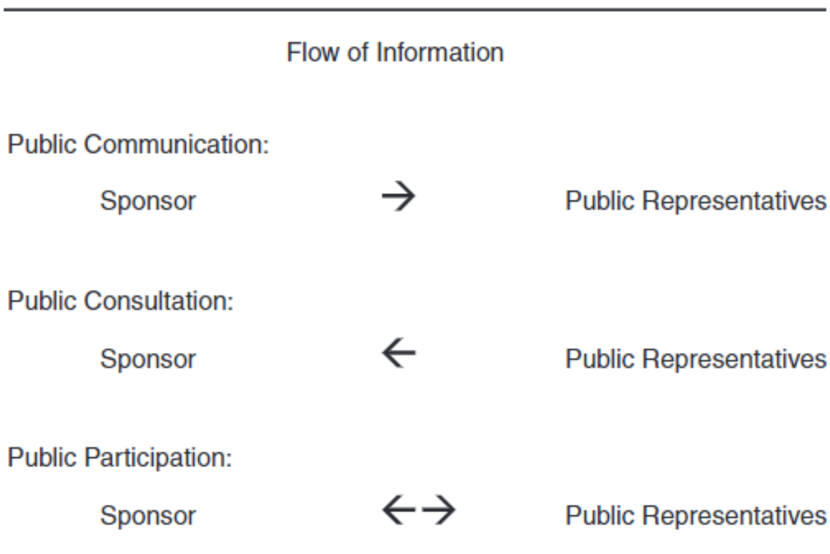
Another example of a way of categorising engagement lies within the Select Committees of the UK Parliament. As their goal is to scrutinize government and put forward recommendations, engaging with the public is a core aspect to their work. Following an internal report, select committees within the UK Parliament were found to fall into one of three categories with respect to their public engagement activities: the traditional, the careful, and

[4] But this could be down to the lack of coordination between departments needed to conform to the same plan. There do however exist individual communications and outreach plans in various departments of parliaments which could be brought together.

the innovator (Liaison Committee, 2015, table 8), the latter being the most proactive in using several online and offline methods to engage with the public. In fact, throughout the whole process of the inquiry stage, select committees were found to alter their approach in engaging with the public depending on the stage of the inquiry they were working on. Gathering evidence for an inquiry is vastly different from letting people know about the launch of the final report, for example. These categories (the traditional, the careful, and the innovator) are a description of the end result of particular activities, but still help to identify the original underlying dimension of engagement.

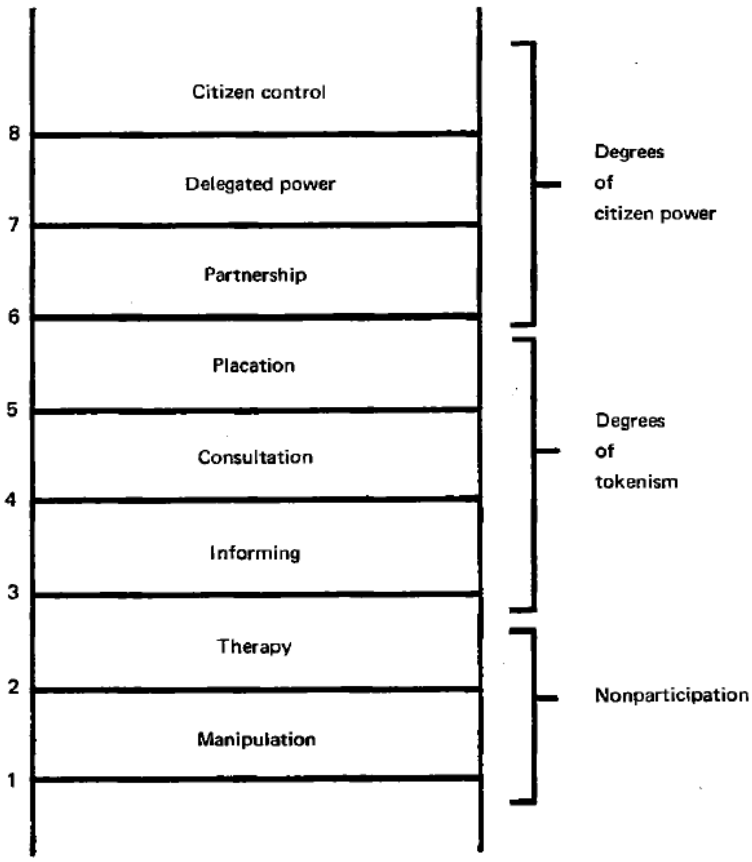
Several authors including Hansard Society (2011); Leston-Bandeira (2007); Walker (2012); Coleman and Gotze (2001); OECD (2009); Leston-Bandeira and Walker (2018) and Kalampokis, Tambouris and Tarabanis (2008), separate the areas of dimensions of engagement into between 2 to 5 categories each focussing on different aspects of information and participation. They follow similar principles whereby the different approaches taken by parliaments broadly fall into one of two camps: those that seek to inform the public, and those that seek for the public to inform the parliament. The first involves a one-way conversation, while the other is much more collaborative on the institution's side and participatory on the public's side. This interpretation of public engagement as a particular flow of information is echoed by both Rowe and Frewer (2005) and Lenihan (2008). Figure 2 from Rowe and Frewer (2005), explains their interpretation of public engagement, which consists of three dimensions: *Public Communication, Public Consultation*, and *Public Participation*. Interestingly, along with the flow of information, the *Public Communication* category does not make any clear assumptions of knowledge on the part of the public. Similarly, Lenihan (2008) focuses on models which involve consultation with the public at a risk of creating distrust – *Consultation Model*, and a model that encourages true conversation and dialogue with the public - the *Public Engagement Model* (p.16-18). The latter's model pre-assumes the public's knowledge of the institution and therefore just focuses on the participatory aspect of engagement. Contrastingly, Rowe and Frewer's model in Figure 2 includes a *Public Communication* category which involves the sponsor being the sole source of information by communicating with the public, a dimension not included in Lenihan's model.

Figure 2: Types of public engagement (Rowe and Frewer 2005, p.225)



However, engagement has not always been interpreted as a flow of information. Generally, the end result of engagement from the public's point of view is to have their voices heard by an institution and to make a difference. Therefore, along with understanding who is initiating the discussion, it is also important to understand how much effect that discussion will have on the institution – or more simply, how much the discussion helps the participants in achieving their end result. Arnstein's seminal study (1969) addresses exactly this: the extent to which the inputs generated through public engagement have an effect on the relevant political institution and/or decision-making. Arnstein (1969) conceptualises engagement in terms of a ladder of public participation, as illustrated in Figure 3. This ladder conceptualisation identifies two polar opposites: the perceived/misleading participatory activities at the bottom and the actual/genuine participatory activities at the top; signifying increased influence over the system as one progresses up the ladder (Arnstein, 1969). This is an intuitive way of thinking, and in many ways comes from the perspective of the public's perceived input of the engagement activities they are undertaking and their actual effect on policy. Furthermore, this emphasis on the public's perception of engagement initiatives is in contrast to the other institution-centred approaches such as Rowe and Frewer (2005). This approach shows that although an institution may be appearing to engage with the public, the amount to which they are truly participating with them may not be equal. Only when the activity results in the institution not only hearing but heeding the views of the citizens, does Arnstein argue that citizens truly have power.

Figure 3: Arnstein's (1969) Ladder of Participation

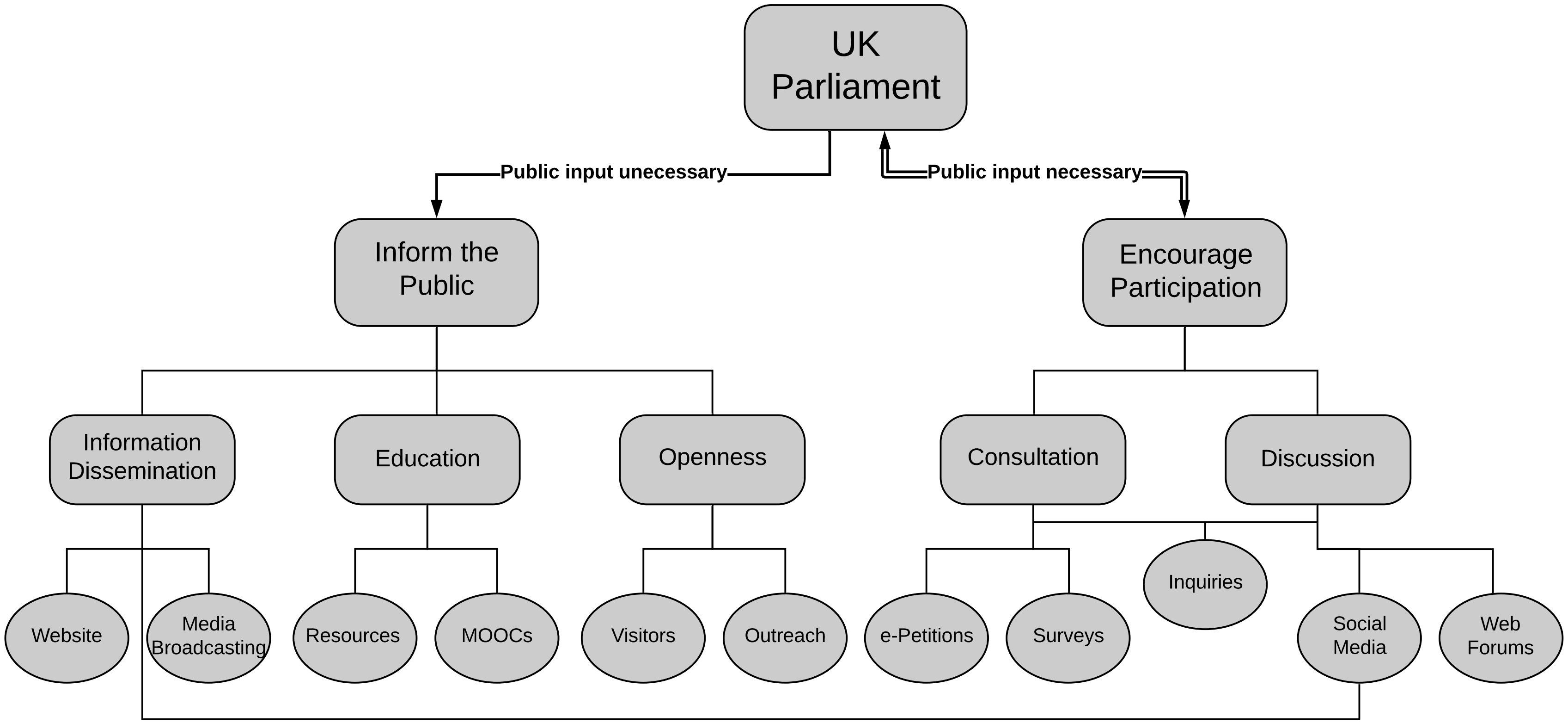


Taking these prior interpretations of public engagement into account, including examining and reflecting on the work currently underway in the UK Parliament, in Figure 4 I propose an alternative interpretation to Figure 2 and Figure 3. The interpretation of public engagement to be taken in this project is based not only on the flow of information, but the source of input. The spectrum model incorporates the different dimensions of engagement (second tier) and sub-categories within these dimensions (third tier) involved in public engagement, as well as providing real-world examples of how the theory can be realised in different institutions (fourth tier).

This process of public engagement is split on the second tier between informing the public and encouraging participation which are at equal levels of the spectrum. Following the left-hand side path involves a one-way conversation out of Parliament, while the right-hand side path involves more mutual communication. The third tier represents the categories of engagement that lie under the two ends of the spectrum, while the fourth tier provides examples of specific outlets that are used by the UK Parliament to achieve the categories above. The examples on the fourth tier are specific to the UK Parliament but can be generalised to other institutions. For instance, many institutions will use tools such as websites and media broadcasting to inform their audience of the latest news. Likewise using ready-established external communities (online or offline) to hold a discussion on a particular topic or to get valuable input is not uncommon in many institutions.

Relating to the models above, the left-hand side can be likened to Rowe and Frewer's *Public Communication (*Figure 2*)*, or Arnstein's lowest rungs of *Therapy/Manipulation (*Figure 3*)* which she categorises as 'Non participation' (Arnstein, 1969, p.217). Conversely, the right-hand side incorporates the *Public Consultation and Participation* categories of Rowe and Frewer, and the higher rungs of Arnstein's ladder. The *Public Consultation* category from Figure 2 is still one-directional but importantly, represents a perspective from the public. This is also echoed by Coleman and Gotze (2001) in that they also make a distinction between information, consultation, and active participation. By providing their views on a certain matter in Parliament as is the case for this category, the public is indeed participating in parliamentary matters beyond simply consuming information. Crucially, without the public's input the left-hand side of Figure 4 would remain the same. Information would still be disseminated (be that through the website, Hansard, or television broadcasting), educational resources would still be provided, and visits and outreach initiatives would still be made available and conducted when required. However, in order for some of the activities in the right-hand side of Figure 4 to be accomplished, the public's input is vital and a leading underlying contributor to the activity.

Figure 4: Spectrum of digital parliamentary engagement



This model's structure enables a clear and simple way of integrating the key dimensions needed to understand public engagement. It focusses on the flow of information between the public and the institution whilst also categorising dimensions and activities for the purpose of evaluation. In the following chapters, I will show how each of the dimensions in the spectrum can be evaluated and measured using natural language processing and data mining techniques. Keeping the model simple results in one that can be scaled up and applied to a range of circumstances and institutions, making it more robust without losing a great deal of accuracy or descriptiveness. The definition of parameters to explore different elements of this spectrum is something that will be addressed in Chapter 3 and Chapter 4.

I now proceed to explore in more depth two key elements of public engagement: informing the public and encouraging participation, as identified through two branches in the Spectrum of Public Engagement (see Figure 4). I provide examples of how each of the five subcategories on the third tier, *Education, Information, Openness, Consultation* and *Discussion,* have been implemented in various ways to provide a comprehensive account of public engagement in the UK Parliament. The differences in direction flow and source of information become clear with Parliament dominating the left-hand side of the spectrum and the public having greatest influence over the right-hand side.

2.3 Inform the public

Informing the public is perhaps the first and most important aspect of public engagement as seen by several authors (Smith, 2009; Walker, 2012; Leston-Bandeira, 2012; Coleman and Gotze, 2001). Its aim is to provide the public with a solid grounding of knowledge before any type of reciprocal participation occurs. It provides the foundation for meaningful discussion, as it is generally difficult to provide your views on a topic you know nothing about (Coleman and Gotze, 2001, p.6). However, in the parliamentary case this is even more important in that people need and have a right to know about their democratic centre, even if they have no initial desire to participate. The provision of information can and has been realised in many ways in the UK Parliament with just under half of the population claiming to know 'a fair amount' about politics (Hansard Society, 2017, p.38). Visitor services, educational resources, improved websites, and television broadcasting have all been used in varying degrees to provide information about Parliament's work to the UK public (Walker, 2012; Hansard Society, 2011; Inter-Parliamentary Union, 2016). The rationale behind investing time and resources into this dimension of engagement is that if one knows enough about a topic, they will be able to see its relevance to their day-to-day lives and therefore be inclined to participate and engage with the institution (Hansard Society, 2011; Walker, 2012).

On the other hand, the recent trajectory of several pirate parties in Europe show the exact opposite. Because of them providing so much information in the form of webcasting all aspects of their meetings live online, they left themselves open to high levels of scrutiny by the public who could now see any conflicts and disagreements that would usually happen behind closed doors. Members of the party became disengaged leading to some people leaving the party and its ultimate disbanding[5] (Epstein, Newhart and Vernon, 2014; Fredriksson, 2015). People do not want to see people disagreeing and arguing with each other, especially those who are in a position of power

[5] This is the case in Germany, however there still exist Pirate Parties elsewhere namely in Iceland and the European Parliament

and influence (Becker *et al.*, 2012; Mutz, 2006). Furthermore, more information does not necessarily equal a more informed population as information overload can lead to frustrated and overwhelmed citizens (Bartlett, 2018). Nevertheless, making the electorate aware of parliamentary business is vital to democracy, but the way in which this is achieved needs careful consideration. Relating back to the spectrum of engagement, this dimension of engagement will incorporate not only information, but the education of the public and the openness of the institution, as per Figure 4.

2.3.1 Dissemination

Providing information to the public on parliamentary business is paramount, but there is a difference between providing information for the sake of it, and providing information with the intention of people being able to read and understand it. The UK Parliament has been recording the minutes of meetings in the form of Hansard Reports since the beginning of the 20th Century (House of Commons Information Office, 2010), but the number of people outside of the institution actually reading them is low (Hansard Society, 2017, p.32-33)[6]. The creation of websites in 1996 (Coleman, 2004, p.3) gave parliaments a fantastic way of reaching as many people as had access to the internet, much more than who could make the journey to Westminster. They were able to provide the public with a link to Parliament where they could provide all the information they felt a public should know about its institution.

However, Parliament already provided resources to specifically give the public the opportunity to obtain information well before the existence of the internet. A dedicated telephone number for people to contact for any parliament-related questions was set up by the House of Commons in 1979 (Walker, 2012) and is still maintained today by the Enquiry Service. This could be seen as an attempt to normalise contact with Parliament in the UK, making it only a phone call away rather than hundreds of miles many of the public would have to travel otherwise. Media broadcasting was officially introduced to both Houses in 1989 and allowed the public to view parliamentary proceedings from the comfort of their own homes (Norton, 2005).

However, depending on the definition adhered to, this end of the spectrum does not entail engagement at all, it is simply broadcasting of information in one way or another. One could argue that it is an unnecessary step but there is also a counter-argument to why engagement matters in the first place including trust, touched on in section 2.1. Therefore, while ensuring that the public is informed is perhaps the least engaging form of engagement, it forms the building blocks and foundation for true participation and an incentive for the public to engage.

2.3.2 Education

Education and information can go hand in hand, and both fall under the public engagement remit of informing the public. For the public to truly appreciate the business of Parliament, they must first understand why certain measures are in place, and what they are trying to achieve. This dimension of engagement is concerned with efforts Parliament is putting in place to ensure the information they are disseminating is actually understood by the public rather than being absentmindedly viewed. This interpretation of education as a form of engagement has been raised

[6] Note: this refers to reading official Hansard Reports. The number of people listening to a debate or meeting is higher. See Hansard Society (2017) *Audit of Political Engagement 14*, London. Available at: https://www.hansardsociety.org.uk/research/audit-of-political-engagement.

by other researchers. For example, Leston-Bandeira (2012) introduces a 5-step engagement model which includes *understanding* as its second step with an emphasis on how citizens engage with the information they have access to. An educated public is an informed public and tackling the issue of citizenship education from a young age is paramount. Authors in the field suggest getting young people to understand and therefore participate in parliamentary matters is a task facing many parliaments around the world (Inter-Parliamentary Union, 2016; Hansard Society, 2017; Smith, 2009; Parycek *et al.*, 2014). Certain measures have been taken in the UK Parliament (both online and offline) to engage and encourage young people to learn more about what Parliament and its Members do day-to-day.

An educational website was launched in 2008 during the first 5-year public engagement strategy which aimed to encourage school aged children to participate with Parliament's work (Walker, 2012). It was interactive and included several games including the popular '*MP for a week'*. This taught students the broad spectrum of work MPs undertake so they can better understand what goes into representing a constituency. The interactive nature of the game is one of the ways Parliament is shown to embrace modern communication techniques, as well as it being available to anyone with access to the internet. The success of the game is shown by its longevity as it is still running almost 10 years later (Parliament Education Service, 2017a; Inter-Parliamentary Union, 2016).

More recently, an Education Centre was opened in July 2015 to cater to visiting school children. This allows young people to learn more about Parliament on-site, as well as breaking down the financial and institutional barriers to visiting through the transport subsidy scheme, which offers transport subsidies for schools outside of the London area (Parliament Education Service, 2017c; Leston-Bandeira and Walker, 2018, p.308). Outside of London, teaching resources and training are available for teachers to use within their classrooms to educate their students. Resources and materials for understanding what Parliament does, how voting works, and how laws are made are all available for teachers to use. Parliament's Education Outreach team also visit schools to conduct workshops and activities, facilitate MPs visiting schools, and even the opportunity to skype the Speaker of the House of Commons (Parliament Education Service, 2017b). By providing these resources, Parliament is allowing the younger generation to understand and make informed choices by the time they reach voting age of 18.

In this way, active measures are put in place to ensure the public have a mechanism to fully understand and digest the wealth of information coming from the institution. Providing teachers with resources to help them educate their pupils ensures that the vital information about Parliament is being taught in the correct way and that it can be standardised across the country. In doing so, students from all over the United Kingdom have access to the same quality of materials and are not disadvantaged because of where they live or financial constraints.

Efforts to engage and educate younger members of society have also been underway in the devolved legislatures within the UK for some time. For example, in 2019 the Scottish Parliament hosted the annual teachers' Modern Studies Association conference in Holyrood (Modern Studies Association, 2019). This event coincided with the 20-year anniversary of the Scottish Parliament and provided a medium for teachers of Modern Studies in Scotland to come together and gain professional learning experience. The European Parliament also have an extensive provision for education engagement (Leston-Bandeira, 2012). Euroscola invites students from the 28 member states to spend a day in the European Parliament offering "an immersive experience in the Chamber of the European Parliament in Strasbourg, allowing high-school students to learn about European integration by experiencing it first hand." (European Parliament, 2020). This process bears similarity to the Education Centre at the House of Commons in the UK. While many of the education engagement initiatives mentioned so far are focused on the younger generation, Estonia has invested into the improvement of digital literacy for their older generation (Savina, 2016; Empirica, 2014). Estonia already have a very well established digital presence regarding their governmental services and first introduced e-voting in 2005 (Tamkivi, 2014). While this form of digital literacy education is not specifically tailored to parliamentary education, understanding how technology works and how it can be used to one's advantage is vital when dealing with the older generation. Ensuring digital literacy skills across the older population lays the groundwork for more specific online engagement activities to be used by all sectors of society and ensures no one is unintentionally excluded by the new form of digital engagement.

2.3.3 Openness

Another key dimension of the spectrum is Openness, something highly relevant in the context of Westminster. Westminster Palace can seem a daunting place, with its gothic architecture, Victorian traditions and the fact visitors used to be called 'strangers' (Coleman, 2004) added to the sometimes impenetrable façade of the institution. Therefore, in order to encourage the public to understand more about the institution, Parliament addressed the perception of a closed institution and introduced new initiatives for the public to visit. This came in the form of a Visitors Centre (Coleman, 2004), and Central Tours office whose aim was to make Parliament more welcoming for visitors as well as facilitating tours with trained tour guides (Walker, 2012, p.274). The number of school children able to visit Parliament also increased, with the introduction of the Education Centre mentioned in the previous section.

Outreach initiatives are an important consequence of public institutions becoming more open. Encouraging the public to visit and be visited by Parliament was important in their perception of the institution. Members of the public already interested or actively engaged in politics most likely already knew about the ways they could visit Parliament, sit in on Select Committee meetings, or Westminster Hall debates. Taking Parliament to the people who were under-represented was a new focus, and can be represented as a characteristic of the *Mediator Parliament* (Leston-Bandeira, 2016). This type of Parliament developed from 2000 onwards and directly corresponds with the timeline in Figure 5 whereby the priorities of Parliament changed from one that provided basic information to the public to one that actively involved the public. This new focus on public engagement was in great part a result of the influential Modernisation Committee report (Modernisation Committee, 2004) which caused Parliament to take the relationship with the public seriously. As Leston-Bandeira argues relating to Parliament's new focus of attention, "It is a shift from a passive assumption to an active role beyond traditional parliamentary business." (Leston-Bandeira, 2016, p.508). The Parliamentary Outreach Service was introduced in 2008 after recommendations from the Administration and Information Committees report. This service has "regional outreach officers engaging with national and local organisations and spreading awareness throughout the UK of the work, processes and role of the institution of Parliament." (Walker, 2012, p.274).

Although an important dimension of public engagement, this dimension largely remains a process of one-way conversation. The public is still receiving information from Parliament with little input going in the other direction. However as several authors argue, knowledge gain (in whichever form) forms the building blocks of successful conversation and participation. Arnstein's lowest rungs of the ladder *Manipulation* and *Therapy* "enable powerholders to 'educate' or 'cure' the participants" (Arnstein, 1969, p.217) and need to be addressed before progressing up the ladder.

Moving away from more traditional methods of engagement to the virtual world, the website was now to include data on the core business of Parliament (Walker, 2012). Some of the measures taken to inform and educate the public also crossed-over into the task to make Parliament more open; for example, the outreach service and webcasting of committee meetings which helped the public view proceedings online, ensuring they were not restricted or excluded by geographical location.

In the previous section, pirate parties were mentioned as an example of when well-intended information-sharing can backfire. They wanted to develop a new version of politics which is open to its members and does not discriminate against hard-to-reach sections of society. Heutlin (2016) states the German Pirate Party ran for "greater government transparency and internet privacy". Transparency is therefore a key theme within these parties, and as such, novel ways by which their members could view and participate with party business were used. For example, the German Pirate Party livestreamed every one of their meetings online so anyone was available to view and keep up-to-date with party business. However, this was a double-edged sword. As Becker *et al.* (2012) argue, this caused the party to appear disorganised. Furthermore, in an interesting turn of events, the members of society usually most removed from and underrepresented by politics became the new 'information elite'. They were mainly those who were regularly at home during the day due to unemployment, disability, or other reasons. At first glance, the same criticism could be given to the UK Parliament's webcasting and television broadcasting of parliamentary business, however not everything is captured on camera. The UK Parliament only record details of public meetings, and detailed meetings of a sensitive nature amongst others are not publicised. Furthermore, the UK Parliament is comprised of many political parties with different viewpoints all working towards the same goal – to hold the government to account. As a result, its lack of strong party-political agenda helps avoid arguments which could undermine the reputation and legitimacy of the institution. The fate of the German Private Party may serve as a warning to future political parties or institutions who attempt to become more open without considering the consequences.

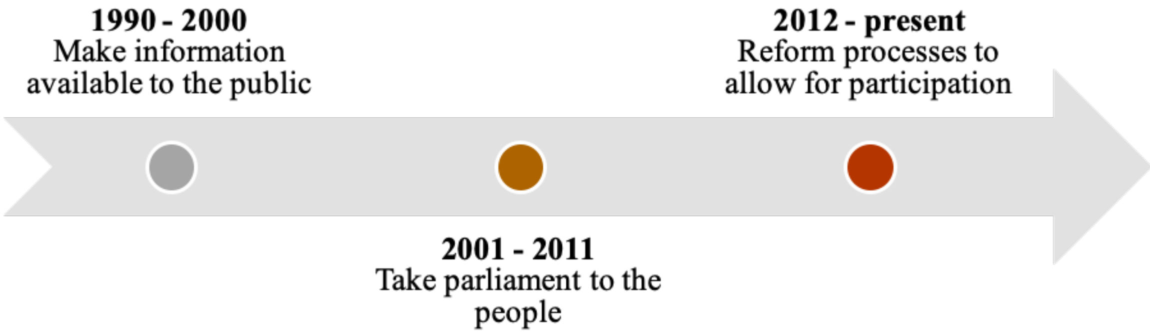
Therefore, addressing openness in Parliament as a dimension of public engagement is not as clear-cut as it may first appear. The degree to which transparency is achieved and the ways it is tackled needs careful consideration, not just for security reasons but for the future reputational consequences as they move across the spectrum of engagement to encouraging participation. Parliament is a naturally very risk averse institution[7] whose purpose is already unclear to the general public (only 46% of the public have knowledge of Parliament (Hansard Society, 2019)). Therefore, doing anything which has the potential to further deter the public from wanting to find out about the work of the Parliament, or lead the public to devalue Parliament would be avoided by officials. This risk-averse nature may contribute to the public's lack of understanding about the institution and feelings of separation.

[7] Discussed further in section 4.3

2.4 Encourage participation

So far, I have focussed on the dimensions of engagement from the spectrum which concern providing information to the public. In these dimensions of information, education, and openness the public's input is not necessarily required for the engagement activities to take place. By contrast, this section concentrates on the dimensions of consultation and discussion and shows how the public's contribution is a vital and motivating factor in these engagement activities. In the case of the UK Parliament, the changing thought process of public engagement is illustrated in Figure 5 below. It can be seen that from 1990-2011 (the end of the first 5-year public engagement strategy), the focus was largely on a one-directional flow of information out of Westminster (Walker, 2012). As the sections above showed, education initiatives and dissemination of information relating to parliamentary business were provided to the public but efforts to actively learn from the public were limited.

Figure 5: Progression of the priorities of public engagement in UK Parliament



However, as shown in Figure 4 and explained in section 2.2 the source of the input changes from the institution with the disseminating information dimensions to the public in the encourage participation dimensions, opening up a clearer avenue for participation. In this right-hand side of the spectrum I focus on activities which enable more debate and collaborative discussion. For example, moving away from tv broadcasting towards petitions and online forums. Furthermore, having the public either as the instigators of the discussion or main contributors, should in theory create a scenario which directly addresses their issues because they are leading the discussion and are able to highlight the problems which are most important to them. Methods of participation which either party is trying to achieve can all lead to the use of different tools or *mechanisms* as characterised in Rowe and Frewer (2005).

Several authors (Coleman and Gotze, 2001; Lenihan, 2008; OECD, 2009) have separated participation into different categories depending on the source of the information; just the public, or both the public and the institution. The former often being referred to as *consultation* and the latter *participation*. However, for the purposes of this research they both fall under the right-side of the spectrum: encouraging participation. The reason behind this is that the conceptualisation in Figure 4 is based partly on how the public's input affects the dimensions. Therefore, as the public's input is vital for both consulting and discussion, they are both classified under the umbrella of encouraging participation.

2.4.1 Consultation

The consultation side of the spectrum involves the institution requiring input from the public. From the literature, the purpose of this is two-fold. First, the institution wants to understand the public's views on a certain topic – perhaps a sensitive subject or from an underrepresented sector of society. Secondly, the institution wants to show the public that it actively pays an interest in their views and experiences. The first purpose can be seen readily in different exercises conducted by the UK Parliament, for example select committee inquiries, public (paper) petitions and e-petitions (Liaison Committee, 2015; Asher, Leston-Bandeira and Spaiser, 2017). The second purpose is down to the way the Parliament receives and listens to the views of the public and how they design their engagement activities to facilitate this.

Arnstein (1969) conceptualisation of engagement as a ladder suggests that many participatory activities do not actually do more than pay lip service to the public. Lenihan (2008) also argues against the *Consultation Model* as a means of true deliberative engagement. He argues it simply causes distrust between members of the public by using shock tactics and exaggerated statistics to out-do their opponents for the attention of the institution. For example, two sides of an argument posted online can call on biased studies or infographics which are intended to catch the reader's attention and discredit the other side rather than truly inform the debate. These posts will then be more overt to the parliamentary officials analysing the consultation session and threaten to derail the whole conversation. Furthermore, the recommendations which surface as a result of the consultation session can be "incompatible" (p.16) with the committee's agenda because they may be too heavily focussed on something outside of the committee's remit. For example, some e-petitions are rejected if they are "about something the UK Government or House of Commons is not responsible for." (Petitions Committee, 2020). This then causes the public to be dissatisfied with the response and perhaps be reluctant to participate again in the future.

An example of this was seen recently in relation to e-petitions on Twitter. One Twitter user who had previously submitted an e-petition had received a response from the Government once it had reached the required 10,000 signatures threshold. Displeased by the response, the user tweeted the official House of Commons Petitions Committee account with the response letter attached (Petitions Committee, 2017a). Despite it clearly stating so on the letter he attached, the Petitions Committee account had to correct him, explaining that his response was from the Government Department for Exiting the EU, not the Committee themselves. However, one could argue the damage had already been done. The Twitter user may have thought that the Petitions Committee was simply paying him lip service as per Arnstein (1969) criticisms. Furthermore, a lack of education or information of the Twitter user may have caused even more annoyance. By not understanding the distinction between a parliamentary select committee and Government department, the user's frustration with the latter impacted on his perception of the former. This relates back to the importance of education and understanding of the institution before tackling participation. In addition, by tweeting the response and showing his anger in public, his followers may also associate participation with the Parliament as a negative experience.

Nonetheless, consultation is a vital step up from Parliament simply informing the public. Understanding the public's views, be it initiating the consultation as with committee inquiries and the process of giving evidence, or when the public initiates the process as with e-petitions, is a way for parliaments to be more engaged with their public. As a counter example to the one above, some e-petitions have had fantastic results and have shown to contribute to Government decisions.

For example, the e-petition on Sugary Drinks Tax was created by well-known British chef Jamie Oliver and the campaign group *Sustain* in September 2015. The petition wanted the government to introduce a tax on drinks containing high levels of sugar. The increased cost of the drinks would in their eyes discourage children and their parents from buying them and encourage them to purchase healthier alternatives. Once received by the Petitions Committee and after having surpassed the 100,000 signature milestone in just 48 hours, the petition was passed onto the Health Select Committee who were leading an inquiry into childhood obesity at the time. The creators of the petition were asked to give evidence to the committee to further inform their inquiry. This led to the committee coordinating their work so the report launch and the e-petition debate in Westminster Hall were held on the same day, which also helped with the inquiry gaining publicity. As a result of the Health Select Committee inquiry (and helped by the e-Petition), the then-Chancellor George Osborne announced in the March 2016 budget that the Government would introduce a levy on soft drinks following a recommendation from the Health Committee's inquiry (Petitions Committee, 2016b; Petitions Committee, 2016a). This tax was officially introduced in 2018. Therefore, in less than 12 months, due to the efforts of the creators of the e-petition and the Petitions Committee facilitating collaboration with another select committee, the Government had changed its mind regarding a tax, and importantly the public who had signed the original e-petition could witness their efforts amounting to a change in the law.

While it is clear that many engage with the Parliament because they want to draw attention to a specific problem, not everyone who thinks to participate with Parliament online does so with the same reasons. A category of "aimless petitioners" were identified by Hale *et al.* (2018) in relation to visitors to the Government's e-petitions website (pre-2015), and found they were highly likely to pick a petition to sign based on what was already trending rather than their own opinions.

As valid as Lenihan (2008) and Arnstein (1969) criticisms may be, namely that participation with the institution is done as a lip-service to the public or that consultation activities can cause distrust and give precedence to the loudest voices, it should not detract from the fact that the public's views are mostly being heard and reflected upon through consultation. However, they do provide an interesting alternative to the motives of institutions holding public engagement activities. What cannot be denied is the single direction flow of information in this setting. The public is indeed providing information to the institution (and therefore participating in a way), however conversation and debate is not generally taking place between Parliament and the people. This raises another question: Is this dimension of engagement about debate **between** public and institution, or debate amongst the public **about** the institution?

Put another way, encouraging participation can have a goal to encourage the public and the institution to engage with each other as in Figure 6. In this scenario, mutual conversations take place between the institution and the public along with a true exchange of ideas. The alternative scenario illustrated in Figure 7 encourages the public to have more of a conversation about the institution amongst themselves. This way, the institution plays a less active role in the conversation and instead it is the public who has a conversation about any topic related to the institution. The latter scenario provides an opportunity for the public to voice their thoughts with each other and then collectively feedback to the institution; but the institution does not necessarily have to close the feedback loop with them. In both scenarios, discussions about the institution are taking place, and therefore being successful from a perspective of boosting the awareness of Parliament's work. However, from a perspective of getting the institution to directly engage in dialogue with the public, it appears only scenario 1 (Figure 6) will do.

This distinction is important because the type of conversation results in different outcomes and being clear from the beginning can avoid any disappointments and manage expectations. For example, for a particular engagement activity, the goal may well be something along the lines of

Figure 6. However, when looking more closely, the actual behaviour of the discussion may resemble Figure 7 more. Understanding this distinction also has repercussions in terms of time commitment and resources of officials vs. keeping a commitment to the public to involve them in proceedings.

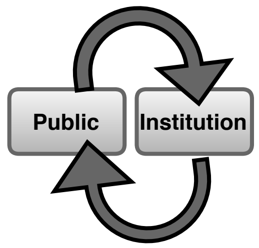


Figure 6: Conversation between public and institution

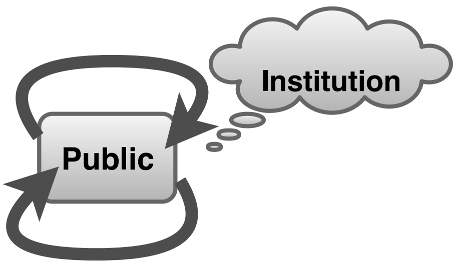


Figure 7: Conversation amongst the public about the institution

2.4.2 Discussion

In this subsection, I examine how parliament engages the public and use examples of several political parties to demonstrate how this dimension of engagement (according to the spectrum in Figure 4) is not specific to parliaments. Rather, this dimension can be (and has been) used in situations with political parties where communicating effectively with citizens is the priority. The question posed above (Is this dimension of engagement about debate **between** public and institution, or debate amongst the public **about** the institution?) makes a distinction between participation for the sake of influencing Parliament or Government (Figure 6), and participation to encourage the public's interest in politics (Figure 7). These are not necessarily mutually exclusive, as the former can entail the latter. Participating in a typically consulting fashion as shown above with petitions can lead to the Government taking a harder stance on a topic, by changing their response or addressing a select committee report. An example lies with an e-petition created in response to a woman sent home from work with no pay for refusing to wear high heels (Petitions Committee, 2017b). In a letter from the Director of the Government Equalities Office, the Petitions Committee were acknowledged for contributing to new guidance for employers developed by the government (Hilary Spencer, 2018). The motivations for both of these scenarios can come from the institution or the public, and as discussion involves a two-way conversation with input from both sides, it has the possibility of encouraging real, meaningful debate.

In Lenihan's report, the Government's role changes from a decision-maker in *The Consultation Model* to a facilitator in *The Public Engagement Model* (Lenihan, 2008, p.16-18): "(i)ts primary task would be to get the various stakeholders to begin engaging one another, rather than competing for influence." (p.17). This in turn would provide the government with a better understanding of the public's views and a more thought-out and realistic set of recommendations, and relates to Figure 7. Going back to the earlier question: *is this dimension of engagement about debate **between** public and institution, or debate amongst the public **about** the institution?*, the public's discussion of issues that are of interest to them can, in turn, raise awareness of these issues within the institution.

The way the institution is set up also appears to be a contributing factor to how parliaments go about participating fully with their public. Although parliaments and political parties are distinct institutions with different priorities and reasons for engaging the public, they can have similar methods to do so. Despite having a rather turbulent history, one area that pirate parties excelled in was communication and engagement of their members. As one of their core messages was that everyone should have their say and a large proportion of the members had a technical background (mainly made up of tech-savvy citizens who were much more comfortable with the online nature of the party), they were perfectly positioned to create their own tools for maximum participation. These tools allowed the party to make true on their promise of mass collaboration with their members, facilitating policy proposals, amendments, and even voting in such a way that favoured their ethos of liquid democracy (Fredriksson, 2015; Heutlin, 2016; Meyer, 2012). This liquid democracy which is at the heart of the pirate parties is defined as "a procedure for collective decision-making that combines direct democratic participation with a flexible account of representation."(Blum and Zuber, 2016, p.165) and "[…] to give each citizen the possibility to vote on each particular issue" (Litvinenko, 2012, p.406).

Likewise, Italy's Five Star Movement champions a model of direct democracy which facilitates online voting and the use of the internet to involve all members of the movement (Mosca, 2014; Natale and Ballatore, 2014). In the most recent Italian elections, the party received 32.7% of the overall votes and formed a coalition government with the Democratic Left political party (BBC News, 2018; New York Times, 2018). However as with German pirate parties, the Five Star Movement has encountered some criticism with its very inclusive nature causing the leader to avoid certain topics of debate such as immigration that would "easily split his extremely diverse electorate." (Mosca, 2014, p.50). This appears to be an unfortunate characteristic of online political participation and bears resemblance to the UK Parliament's own stance on digital debates, where certain topics are actively avoided due to their divisive nature. Nevertheless, the Five Star Movement's anti-establishment and populist nature led to the rise of new media in Italian politics, and their blog *beppegrillo.it* allowed for political debate and conversation amongst people who may not have engaged otherwise.

It is clear that the relationship between political parties and parliaments with the public is very different in that political parties have their own agendas and can promote their political ideas, whereas parliaments, as institutions, must remain neutral and reflect the views of many different political parties at once, all of this represented within parliament. Despite this, the underlying theme of the examples above (of pirate parties and Italy's Five Star Movement) is that they have been shaped by strong online activism and have encouraged citizens to be actively involved in policy making from the beginning rather than as a reconsideration of their priorities. If I relate this back to the interpretation of the spectrum of engagement, only participation at the right-most edge of the spectrum (Figure 4) satisfies their principles and is in-fact a core motivation for parties with this type of ethos, rather than an after-thought. However, as seen in earlier sections their tactics including strong transparency can also lead to disengagement and frustration within the party members.

Despite their differing methods of discussion with the public, the UK Parliament also recognises the importance of true deliberative participation. A new shift to participatory democracy is underway which aims to "provide opportunities for individuals to participate in decision-making in their every-day lives as well as in the wider political system" (Pateman, 2012, p.10). The increased use of the internet and digital tools also provides a catalyst for this participatory behaviour, as the public have a multitude of avenues available to them. They no longer need to wait to be asked for their opinion. Parliament have embraced this new way of working with their increased use of social media and introduction of digital debates in 2015 (Parliament.UK, 2017; Leston-Bandeira and Walker, 2018). This change in how the UK Parliament engages into discussion with the public still needs to be carefully considered to see whether the goal of these discussions is to be fully involved (Figure 6), or just as a bystander (Figure 7). Understanding exactly what type of discussion and level of involvement the Parliament is prepared to have with the public not only helps to manage the expectations of the participants, but also helps the UK Parliament to accurately anticipate and evaluate what type of responses and engagement they are likely to receive from the public.

However, as the evaluation strategies for these are still in their infancy, one can look to other institutions for guidance. As has been established in this chapter so far, despite the different models and conceptualisations, the overall structure of public engagement appears to be universal in that at the most basic there is simply disseminating information, while at its most thorough is a fully mutual discussion between institution and public. That being said, the evaluation of public engagement should in theory follow a similar path. As this research is primarily concerned with engagement in the online sphere and its evaluation, the next section will briefly explore this evaluation with respect to online engagement platforms.

2.5 Evaluation measures of public engagement

Having identified the different elements of public engagement, I now turn to the issue of its assessment and evaluation. Measuring the effectiveness of the particular activity is largely down to what it was trying to achieve in the first place. The evaluation of engagement will clearly be different according to the type of activity it focuses on. The assessment strategy for an education or dissemination activity should and will be vastly different to the evaluation of a discussion activity, for example. Descriptive statistics may well be sufficient to capture the reach of a particular briefing or publication on the Parliament.uk website, but how to measure the types of impact and reflections of the people absorbing this information is another matter. The raw number of likes, shares, and views of a digital tool which has been the main way the UK Parliament has evaluated their digital engagement activities, will only tell half of the story.

Focussing first on the *Inform the public* end of the spectrum, specifically *Information*, many parliaments around the world appear to use descriptive metrics such as those mentioned above. The World eParliament Report (Inter-Parliamentary Union, 2016) provides statistics on the number of parliaments using the internet for engagement, but it does this mainly through descriptive statistics. Nevertheless, while 71% of parliaments report their committees used a website to disseminate information and publish reports compared to 13% using social media, and 74% of parliaments find digital tools an important means to inform citizens about policy, only 34% use websites to gain evidence/comments from citizens and 13% use social media (Inter-Parliamentary Union, 2016, table 36). This shows a large discrepancy between the uses of the internet for different types of engagement – namely dissemination of information vs. consultations; and the perceived benefit of the internet and its actual use. The method of reporting of the IPU also indicates a simplistic way of evaluating the impact of these activities. As of 2018, this discrepancy has reduced slightly with 50% of parliaments reporting their committees use the website, and 20%

using social media to seek submissions from the public. Overall, 76% of parliaments have increased their use of digital in this 2-year period but the way they use digital tools has changed slightly. Their use of websites has decreased slightly between 2016 and 2018, but social media usage has increased with 9% more parliaments using it to disseminate information and 7% more using it to collect opinions (Inter-Parliamentary Union, 2018). Parliaments still appear to use the internet heavily for institution-led engagement (left-hand of spectrum) but very little for public-led engagement (right-hand of spectrum), and this emphasises both that there is a distinction between forms of engagement, and how they are put into action and parliaments' own strategies and/or levels of depth in the way they use digital tools. This will inevitably affect their modes of evaluation.

Researchers in Turkey created and tested a framework using a variation of the SWOT (Strengths, Weaknesses, Opportunities and Threats) analysis using questionnaire and survey data from e-government service users (Osman *et al.*, 2014). The authors used a cost-benefit, risk-opportunity model (COBRA), and focussed on citizens as the key stakeholders of e-government services. This allowed them to quantitatively analyse the data based on responses and use Principal Component and Factor Analysis to uncover clusters based on variability in the data. These clusters were able to be categorized into the four components of the COBRA model, allowing the authors to draw conclusions based on pre-defined hypotheses. In this case, while being able to draw conclusions from their data, the authors are evaluating all aspects of e-government, not specific to engagement. Nevertheless, their focus on citizens and the importance of user satisfaction rather than user benefit (Osman *et al.*, 2014, p.253) has similarities to some goals of public engagement – it is not only important to consider the public benefit from an engagement activity, their levels of satisfaction are also key. This ultimately provides the public with an incentive to re-use the services and continue to engage in the future. The traditionally business-centred approach provides a different angle to the evaluation of internet-based activities, but as it is primarily concerned with institution-led engagement which did not include the evaluation of many comments, the thorough analysis of text data is absent.

Surveys such as the Hansard Society Audit of Political Engagement or Eurobarometer capture large volumes of data about public perception. They provide longitudinal data which allows for a time series analysis of the attitudes of public in relation to institutions. However, these methods of evaluation using surveys can introduce unintentional bias in the wording of the questions which lead to different interpretations by the survey participants. Moreover, gaps in data collection can make analysis and comparisons with previous surveys more difficult. Nevertheless, they provide a valuable insight into the public's own opinions and feelings about different democratic issues. The results from these surveys can be compared with and enhance more detailed analysis of specific participation platforms as will be explained below.

Moving now to how the literature evaluates the *encourage participation* side of Figure 4, political institutions around the world use many different online participation tools and have led to the creation of several evaluation frameworks. For example, when analysing how citizens participate in government consultations in Denmark and the UK, Rasmussen (2015) argues that the initial design of a consultation session directly influences its results. One example lies with the *OurSpace* platform which was developed by the EU to encourage young people to be more involved with politics. The project involved several EU countries and was live between 2012 and 2014, with 18 Members of the European Parliament (MEPs) participating (Parycek *et al.*, 2014). The creators of the platform used a 4-point evaluation system measuring four categories: political, social, technical, and methodological, mixing qualitative and quantitative methods (Parycek *et al.*, 2014, pp. 4-5). The categories are as follows: “aspects of influence on political decision-making and the relevance of the discussions for politicians" (political), “aspects of society related to community-orientation and connection” measuring “community building, digital connections between users, and integration of multiple communication channels” (social), “Assessment of platform and tools usability and suitability.” (technical), and "effectiveness of the essential success factors and characteristics of the platform ” (methodological). Approaching the evaluation in this way allows for different areas of the platform to be measured independently of each other, in order to understand precisely which areas are performing well or not. Various methods of evaluation were used including questionnaires, discourse analysis, and interviews both with people using the platform and industry experts.

In the technical category, data was captured to show the number of visitors, and comments and revealing a fair number of people lurking on the platform without posting a comment. They also found many users doing very minimal participation such as liking each other’s comments which required little effort. However, users did rate the platform highly in this category due to its usability. Socially, the consensus was mixed with regards to how much more involved youngsters were in politics or their levels of trust. Furthermore, the six different languages supported in the platform provided an added difficulty in the general evaluation. The UK Parliament only works in English when engaging with the public on social media so the evaluation of different languages was not a hinderance to this thesis.

Finally, politically, they found controversial topics to be most popular but minority rights, education, and environment also attracted good discussions. Ultimately, they found it difficult to involve and “engage young people beyond the already politically interested via an online tool” (Parycek *et al.*, 2014, p.11), which bears resemblance to other studies including Hansard Society (2017); Inter-Parliamentary Union (2016) but calls into question the effort the UK Parliament and other parliaments put into engaging with the younger generation as mentioned in section 2.3.2 . On the other hand, they also noted the presence of the MEPs in the discussion boards and the use of simple language were paramount to the higher levels of engagement and success of the platform. Therefore, in the case of *OurSpace*, using a wide range of evaluation methodologies allowed for a well-rounded approach where different aspects of the usability of the chosen online platform and what influence public officials can have on how participants approach the engagement activity is measured.

Another example of an evaluation measure lies with Macintosh and Whyte (2008) who developed a framework of evaluation for four local government-led engagement activities in the UK. The authors used three sets of criteria with details on what the activities should achieve with respect to the project, democratic, and socio-technical aspects. A range of quantitative and qualitative methods were used including interviews and questionnaires, with web analytics also being used to evaluate the success of the activities with respect to the criteria. Through this approach, as with *OurSpace*, they were able to draw some conclusions about the effectiveness of the four local-government projects and able to put forward their feedback and recommendations to the local authorities involved. The collaboration between researchers and local government already makes the evaluation framework impactful as it shows the practical implications of their evaluation framework, however the degree to which true effectiveness can be measured in this way is debatable even among the authors. They state, “We should note however that there are no standard definitions of effectiveness in eParticipation, nor should we expect any to emerge.” (Macintosh and Whyte, 2008, p.11). Therefore, they manage their expectations in terms of the promises they can make on behalf of the evaluation platform because of the underlying problem facing public engagement thus far that there is no single definition of engagement. Consequently, there appears to be a high level of subjectivity in monitoring the projects mentioned in this section, especially when the input of citizens is considered in Consultation and Discussion activities.

Staying with the valuation of local government, there has also been research into the evaluation of citizen participation using different channels of social media. Spanish researchers evaluated the links between follower numbers on local government Facebook and Twitter pages and levels of citizen engagement. They used metrics of popularity, commitment, virality and engagement. These measured the mean likes per each post (popularity), mean number of public comments on each post (commitment), mean number of shares for each post (virality), and a sum of the first three metrics for overall engagement. They found that Facebook was generally more successful at engaging the public than Twitter but a high numbers of followers did not correlate with high levels of engagement in terms of likes, comments, and shares (Haro-de-Rosario, Sáez-Martín and del Carmen Caba-Pérez, 2018). They also evaluated engagement based on how transparent the local government organising the account was (based on the Spanish transparency indices), and whether the comments were positive or negative. This study provides insight into the use of social media in the public sector and its influence over how the public engages with the public, but does not make a clear distinction between the different types of engagement.

The examples above were mainly concerned with evaluating existing public engagement, specifically e-participation projects. This can only be done once the underlying dimension of engagement which the project is targeting has been decided. The outcome of the evaluation may be changed depending on the particular dimension which the engagement activity falls into. An example of a participation model which factors in the stakeholder, participation process, and details of the ICT tool used, is found in Kalampokis, Tambouris and Tarabanis (2008). A sub-domain of their model is shown in Figure 8 and includes five different participation levels *E-Informing, E-Consulting, E-Involving, E-Collaborating,* and *E-Empowerment,* – some of which mirror the spectrum in Figure 4 and the eight levels proposed by Arnstein (1969). This not only reaffirms the earlier argument of different conceptualisations having a common pattern in terms of the dimensions of engagement, but also provides an alternative evaluation method to other measures (e.g. (Inter-Parliamentary Union, 2016; Osman *et al.*, 2014; Parycek *et al.*, 2014)) which incorporates the grounding theory. The inclusion of the different stages in policy cycle and specific areas of participation make it a holistic and comprehensive model, and one that in theory could apply to various institutions dealing with the e-participation domain of public engagement. However, the extent to which it can be applied to the data types that will be encountered in this research – namely text data, is unclear.

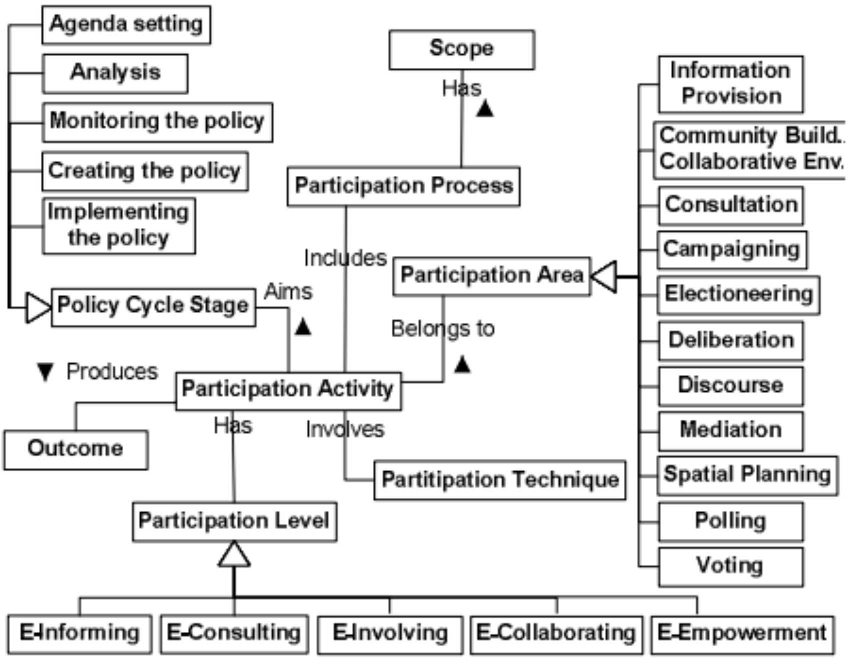


Figure 8: UML Class Diagram of the Participation process model
(Kalampokis, Tambouris and Tarabanis 2008, p.28)

Evaluating the digital dimension of activity also adds another dimension of people's attitudes online which need to be taken into account. Many praise the use of online tools to facilitate engagement, however they must be careful not to exacerbate any offline characteristics. The divide between those who are represented and those who are not is wide, and using digital tools to bridge that gap may not always be the easiest solution. Coleman and Gotze (2001) describe *technopopulism* as "…whereby the loudest, best resourced, most confident or most prejudiced voices of the public come to dominate the debate" (p.8), something that can be seen in offline group discussions and focus groups (Krueger and Casey, 2014; Stewart and Shamdasani, 2014). However, in a focus group setting, the moderator can usually recognise and address the participants' behaviour. That may not be as easy online, however people can be found also to be much more vocal and daring online than they would be in person, be that due to their interpretation and influence of their avatar or sense of *extended self* (Belk, 2016; Yee, Bailenson and Ducheneaut, 2009). Furthermore, digital tools may in fact emphasise the discrepancy in representation leading to a situation where "groups that are traditionally marginalized in political discourse are more likely to be marginalized in the online political discourse as well" (Epstein, Newhart and Vernon, 2014, p.339). This can be a disadvantage when it comes to 'trolls' which could be seen as a potential disruption to online engagement, but could also mean those who would generally be very reserved feel that they can have a louder voice.

While potential difficulties of digital engagement are something to take into consideration with the use of any online engagement tool, using digital tools is something that can greatly help improve the effectiveness of participation and to increase trust in the institution (Warren, Sulaiman and Jaafar, 2014). Specifically during 2020 where the covid-19 pandemic restricted many offline participation activities, having the processes and tools in place to successfully communicate and engage online with the public means that engagement sessions remain of value to both the public and institution. Assessing and evaluating public engagement is just as important as conceptualising the underlying dimensions, and as this chapter has introduced, there is no one rule. Ultimately, the awareness of these dimensions can directly influence the methods of evaluation which are taken. One-directional information flow out of the institution as described in section 2.3 may simply require a more summarising approach where core metrics describing the reach of the particular activity may be sufficient such as those defined in Haro-de-Rosario, Sáez-Martín and del Carmen Caba-Pérez (2018) or the Hansard Society audits of political engagement. However, when the public become the initiators or core contributors to an activity the importance of evaluating in a way that incorporates the range of information provided by the data (for example the linguistic patterns encoded in text, or the difference between participants' responses to various topics) is crucial. These more detailed methods of analysis are not restricted to the evaluation of participatory online engagement activities (as illustrated by the right-hand side of the spectrum) and can also be used for activities which are more concerned with providing information to the public. However, this distinction between methods of analysis and evaluation for the two ends of the spectrum is dependent on the main aim of the institution when carrying out the activities. On the left, their primary aim is often to inform and /or provide resources to the public. On the right, the aim is to gather opinions and information from the public. As a result, the baseline evaluation metrics for the two types of engagement should be different.

2.6 Evaluating digital engagement: an example from select committees in the UK Parliament

This sub-section serves as an example and how public engagement is implemented by a specific service in Parliament. This also illustrates the need for evaluation and the difficulties arising from the type and volume of material submitted. Within the UK Parliament, there are a range of departments and teams which have a remit over public engagement, some of which have already been mentioned. However, one area of Parliament which is particularly concerned with public engagement are select committees. Departmental select committees were first introduced in 1979 and serve to scrutinise the work of government departments (Norton, 2005). In 2012, public engagement became one of their core roles (Walker *et al.*, 2019). They conduct inquiries based on important issues in their jurisdiction and consult the public at every stage of the inquiry process, however research suggests some committees do this better than others. Committees differ in terms of what kind of engagement they carry out. Some, categorised in Liaison Committee (2015) as the 'Innovators', are more comfortable using digital tools to communicate with the public, while others, the 'Traditional' or 'Careful', prefer a more traditional offline approach to engagement. Many also adhere to the 'go where the people are' ethos of engagement whether that be travelling to local shopping centres or accessing pre-existing online communities (Liaison Committee, 2015).

It is understood that building public engagement in politics can be especially difficult partly due to the public's perception of what politics is. A report by the Liaison Committee noted that "Democracy is 'good' but 'politics', or at least overly partisan politics, is bad" (Liaison Committee, 2015, p. 12). The inherently unpartisan nature of select committees which bring together cross-party Members of Parliament to scrutinise the government makes them uniquely placed to tackle this perception of party politics and creates an environment for issue-based engagement rather than party-based. With this in mind, over the years select committees have used digital engagement in various ways during their inquiries.

The different stages of an inquiry also provide select committees with different opportunities for various types of engagement with the public. Agenda setting is used to explore topics before officially deciding to launch an enquiry. These topics can be based on the agenda of the department the committee is shadowing. For example, at the beginning of an inquiry, the Transport Committee used YouTube to set the agenda and invite people to suggest topics, while the Education Committee used a web forum, The Student Room, to ask for opinions on their inquiry into services for young people (Education Committee, 2011). During an inquiry, the social media platform Twitter is often used by several committees to help in the gathering evidence stage. This involves asking the public to comment or ask questions using specific hashtags (for example #AskGove by the Education Committee (Education Committee, 2012)). Some evidence sessions are also live tweeted or uploaded to YouTube to allow the public to view and comment. This is often the case with the Petitions Committee who frequently provide a hashtag for their Westminster Hall debates of e-Petitions which reach the 100,000 signatures threshold. Twitter has seen the biggest increase of use in this stage of the inquiry process, as well as various web forums such as Mumsnet (House of Commons, 2019b), The Student Room (The Student Room, 2020), Money Saving Expert, and Reddit.

However, by embracing digital technologies in their work, certain committees have also found that their internal processes are not equipped to cope with the increased volume of data which is created. The Political and Constitutional Affairs Committee (PCAC) expressed concern for a future ‘engagement explosion’ after receiving 16,000 responses to a voter engagement inquiry. This is recognised as a potential internal barrier to engagement (Nesta, 2019) which can be ameliorated by using digital analysis tools such as TheGist application which condenses and summarises large volumes of textual data from social media and web forum comments to provide a clear overview of a discussion. This application is described further in Chapter 7.

Furthermore, while social media has proven a useful channel of engagement for certain stages of the inquiry process, it has not been used by select committees to select witnesses for evidence sessions. This suggests that committees prefer to seek out established organisations in the more traditional offline manner for help in obtaining witnesses from more diverse backgrounds. Social media has also primarily been used by committees in “broadcasting-out than on seeking input and views” (Liaison Committee, 2015, p. 47). This can be directly mapped to the spectrum of engagement with the majority of digital engagement tasks done by committees falling into the left-hand branch of *Inform the Public* which encourages a one-way conversation and a one-directional flow of information out of the institution. This suggests a need for more deliberative engagement characterised by the right-hand branch of *Encourage Participation*.

Conclusion

What does an engaged Parliament look like? Understanding the characteristics of an engaged Parliament depends on many factors. This chapter has focused on the different definitions, dimensions, and categories of public engagement with some real-world examples from the UK Parliament including a closer look at select committees, and further afield. Understanding what public engagement entails at its different levels allows researchers and practitioners to develop appropriate methods of evaluation. Although many different conceptualisations exist, by analysing the literature many authors are in agreement with the general processes of engagement, but some question the effectiveness of certain methods such as consultation and the real motivations behind the institution's activities. I propose a simplified model of public engagement (Figure 4) which separates activities based on the source of the input, i.e. solely the institution or the public, as well as the actual requirement of the public's input at all based on the original aims of the institution in creating the engagement opportunity. Addressing all aspects of the spectrum (Figure 4), flows of direction, and sources of input should result in a Parliament which is not only shown to care about the views of its public, but provides adequate resources to educate and inform its public too. Tackling only one half of the spectrum could result in a passive public or one that is frustrated with the lack of transparency of its Parliament.

The next chapter focusses on the methodological tools used in this thesis and how the evaluation of public engagement activities at either ends of the spectrum can be achieved and made better through these methods.

Chapter 3 Methodological framework for data analysis

This chapter explains the quantitative data analysis methods to be employed in this project. These methods fall under two areas of research: text mining and social network analysis. The first will look at decoding the underlying information encoded into the text comments of the participants of certain online engagement activities, and understanding at scale what the general feelings of the public are. The second technique will look less at the individual words, and more at the overarching network of the participants on social media – a method of analysis which can be useful in an online setting. Chapter 5 and Chapter 6 will put the methods and algorithms covered into practice using data from online engagement activities run by the UK House of Commons.

As Maynard *et al.* (2017) and Faralli, Stilo and Velardi (2017) show, most social media analysis is undertaken with the end goal of either recommendation of certain products or classification of users for predictions of behaviour (for example voting). This project is not directly concerned with either of these goals. Its aim is primarily to understand what is happening in terms of online parliamentary public engagement in the UK and to develop methods and tools to evaluate this, and so will take a more exploratory approach rather than one focussed on predictions. Elements of classification are used in terms of sentiment analysis and topic models, however these will be to inform the assessment of effectiveness of certain engagement activities. The knowledge gain from the use of the types of algorithms covered in this chapter will be able to inform future decisions on all areas which have an element of digital engagement.

While the majority of the methods in this research involves quantitative data analysis, I also conducted participant observation through regular visits to Westminster. The purpose of these visits was to work with the Digital Engagement team and various select committees to understand how they run their online engagement sessions and how the research can help them to solve problems they face with online engagement. As a result of these visits, I learned many details about the internal processes of parliamentary engagement and how daily business is impacted by internal structures and barriers to engagement. Over the three years, I have had many informal discussions with parliamentary officials concerning how they view their role in engaging the public and what difficulties they face or frustrations they have. This inside view gives me a better understanding of the reality of working in a public facing institution, while being clear on any methods I use or suggest being practical and achievable within the parliamentary context. I was also able to access exclusive data from various UK House of Commons social media accounts and select committee formal evidence submissions as a result of the collaboration and this provided me with a unique insight into parliamentary engagement activities and its outcomes.

The chapter will be structured as follows; section 3.1 will outline the data sources to be used in the research; section 3.2 will introduce the text mining methods including the pre-processing stages to be undertaken prior to any further analysis, as well as the topic modelling algorithms to be used. Further details about the R Shiny web application “TheGist” are explained in detail later in the thesis; and section 3.4 will introduce the network analysis methods and how online communities can be identified as well as the characteristics they may hold. Finally section 3.5 covers the participant observation techniques I used during my visits to Westminster.

3.1 Data sources

The introduction of social media has created excellent avenues for research, with insights into disaster relief or political issues (Verma et al., 2011; Colleoni, Rozza and Arvidsson, 2014). Social media has been a beneficial data source for both academics and those working in industry. For example, a search for 'social media' on Google Scholar produces over 4 million results where academic researchers from multiple disciplines have published studies. Likewise, to rationalise their use of social media in one of their projects, the British Cabinet Office remarked that social media "…is often far quicker and cheaper than other forms of analysis and data is available in or close to real time." (Social media research group, 2016, p. 7). Along with real time collection of data and opinions from the general public, the frequency of posts, timelines, and locations of users can be extracted from social media data. These variables can greatly enhance the interpretability of the results of any analysis. For example, instead of simply having the raw number of posts or tweets in a given timeframe or for a given topic, in certain platforms, I can also see where in the country the users are coming from and the time of day they are most active. This supplementary information can be very valuable for researchers or analysts who require more details from their data and is often extracted using an API. An API (Application Programming Interface) facilitates the communication between different applications. They are primarily used to extract data from an application or website, and allow an analyst to specify what type of data they would like and in what format. I have used the Twitter Streaming API to extract real-time data, in particular tweets, that are added to a dataset in real-time (Russell, 2011; Social media research group, 2016).

Another advantage of using social media data is the level of information relative to the time and effort needed from traditional research methods such as face-to-face interviews and ethnography. Edwards *et al.* (2013) categorise social media research as extensive and real-time as opposed to intensive social research methods such as ethnography and time-bound methods such as surveys and experiments. The distinction between real-time and time-bound separates research that captures social relations as they happen versus ones that captures social relations at a particular snapshot in time. A method becomes intensive when it reflects the views of a smaller set of people rather than an extensive method which has the ability to capture population-level data. Although social media research is described as extensive in this context, there is an argument for it also being categorised as intensive, because specific communities and groups of the population which would traditionally be accessed through ethnography and interviews could also be reached through their online presence (if one exists). An example of this is in the House of Commons Web and Publication Unit's work with a small community of fisherman in the north of Scotland. They used an online Facebook group to communicate with this community of 400 to gain their views on an upcoming select committee inquiry. This shows that while social media can often be successfully used to research population-level topics by extracting information from users around the world, it can also be used to target specific smaller communities.

However, use of this data source also comes with its own pitfalls. Social media platforms are increasingly limiting the volume and quality of data that can be extracted, resulting in fewer types of analysis being carried out (Burgess and Bruns, 2012). A disadvantage of social media data is that it is typically very unstructured and of a low quality in terms of consistency of data points. For example, public tweets can be extracted from Twitter using their API, however the location and biographical data from the author of a tweet is given through a free text input (rather than a pre-defined list) and may not be an accurate depiction of themselves. For this reason, getting accurate information, for example plotting locations of Twitter users, can be difficult. Unlike traditional social research methods such as focus groups or interviews, it is difficult to know who is participating online without explicitly asking for this demographic information. There are different ways to circumvent this restriction of social media data and infer certain demographics from users which will be discussed further in section 3.2.4.

Furthermore, although social media can be used to study population-level data, social media data is not representative of the wider population. There are 13 million Twitter users in the UK (Statista, 2019b) which equates to only 20% of the total UK population (Office for National Statistics, 2019). However latest figures suggest there are up to 44.8 million UK residents using Facebook (Statista, 2019a) which is 66% of the total UK population. This is only a few percentage points lower than the most recent 2019 UK General Election turnout of 67.3% (House of Commons Library, 2020), therefore calls into question the lack of veracity in the representativeness of Facebook data. The majority of the data used in this research will be comments from Facebook digital discussions and tweets using a specific hashtag from parliament's Twitter accounts. The Twitter REST API (Twitter, 2019) is used to extract tweets while the Socialfy platform is used for Facebook comments (Socialfy.pw, 2020). Public tweets can be extracted by anyone who has a Twitter developer account, and I used the Tweepy python library (Tweepy, 2020) to access the Twitter Streaming API. This allows me to search for a particular keyword or hashtag and returns all tweets using the search term in real-time. The python library returns a JSON file with 50 variables, however for the purposes of the analysis in this thesis I only retained six of these variables; Time/date, text of the tweet, user screen name, number of followers, number of friends and user location. The first two variables were used to plot the time and date of tweets, and to create a corpus of tweets from the text of the tweet. The user location is used to plot the geographical area of each tweet, however as mentioned earlier this is self-reported by the user and is often inaccurate. The number of followers and friends in the dataset is used along with the geographical area to determine locations which had the highest number of followers and friends per user. I was selective with the number of variables I ultimately used in my research because the UK Parliament can be very risk-averse when dealing with social media data and did not want to extract more data than was necessary. The Twitter data from the API was primarily used for the analysis in Chapter 6 where I focussed on hashtags for different digital debates and e-petition debates.

A benefit of the collaboration with the House of Commons is the extra access to datasets that I would not have had otherwise. One of these datasets is the list of followers and account-level information about these followers for 48 Twitter accounts owned by the UK Parliament. The majority of these accounts are owned by different select committees while the others are spread over various departments in the House of Lords and House of Commons (see Table 3). This was a large dataset containing almost 2 million records bought by the Web and Publication Unit in 2018 so they could get a clearer understanding of who follows them on Twitter. A detailed analysis of this dataset is covered in sections 5.1 and 5.2.

Thanks to the collaboration with the House of Commons, I was assigned as an analyst on the UK House of Commons Facebook account allowing me to access analytic information from the page, as well as use a third-party application called Socialfy to extract the comment data for further analysis. The possible reach for Facebook is vast but the UK House of Commons official page has an average daily reach of 21,000 users, and a total of 60,000 followers which greatly reduces the overall reach and reduces the representativeness of the data. Facebook also provides certain demographic information within the 'Insights' section of a page. This also includes data on various aspects of a page that an owner could use to assess its performance, i.e. the number of page views, page likes, or user engagement with posts. According to these page analytics, this account is followed primarily by males aged 25-34 and 35-44 age brackets. I extracted the comments from individual posts which were used as digital discussion cards where the public could comment on specific questions raised by MPs. These comments were downloaded through the Socialfy platform and included the date of the comment, the text, and the name of the user who posted the comment. As with the Twitter data, I selected only the time, the text of the comment as the teams I was collaborating with did not want to use any personal identifiers. These restrictions on the type of data and analysis I could perform as a result of the collaboration are raised further in section 3.5.

The final major dataset used in this thesis was from the online discussion platform Discourse. This came from demonstration tests completed with three select committees in 2019 to assess how the platform of engagement changed public participation and opinions within certain topics. The use of Discourse platform was in part chosen due to the ease of data extraction following the discussions. The platform has a data explorer which allows owners of a forum to query the database using SQL queries. Along with the usual statistics for analysing platforms similar to those in Twitter and Facebook, I was able to write specific SQL queries and export the text comments, IDs, and IP addresses of each user, an edges table of user interactions with each other for social network analysis. The results of this analysis are explained in detail in section 6.6.

The majority of the data covered in this section relates to online social media data, but I also made use of data which did not come from an online source. The Web and Publications Unit provided me with a list of individuals and organisation who had made a formal evidence submission to a select committee between May 2016 and May 2018. This dataset contained over 30,000 anonymised records each with details of the committee and inquiry name, submission ID, and the submitter organisation name (anonymised in the case of individuals). This dataset was used to explore which committees had the greatest number of submissions and I used social network analysis to see how the committees and inquiries compared in terms of shared submitters. Further details can be found in section 5.3.4.

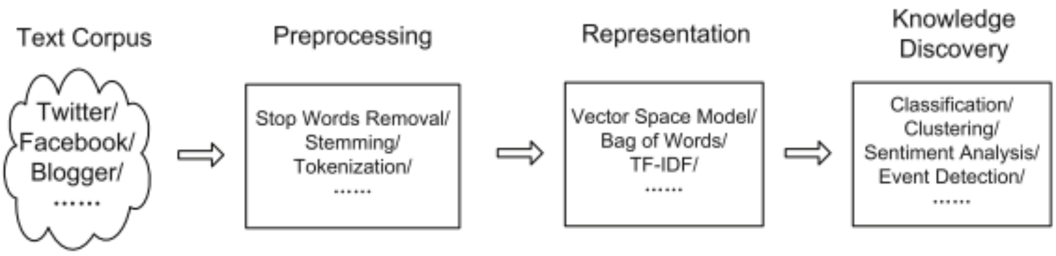
Table 1: Summary of data sources and location in thesis

Dataset	Source	Discussion in thesis
Digital debates	Facebook (Socialfy.pw)	Sections 6.1, 6.2, 6.3, 6.4
e-Petitions	Twitter API	Sections 6.1, 6.2
Facebook user network	Facebook (Netvizz)	Section 5.3.3
Twitter account followers	Web & Publication Unit (UK Parliament)	Sections 5.1, 5.2, 5.3
Demonstration tests	Discourse	Sections 6.5, 6.6
Select committee evidence	Web & Publication Unit (UK Parliament)	Section 5.3.4

3.2 Text mining

As explained in the previous section, this thesis focusses on the analysis of social media and primarily the comments left by participants of online engagement activities. Figure 9 (Aggarwal and Zhai, 2012, p.389) outlines an interpretation of the framework used to analyse text data from social media. They show the different stages in text mining from pre-processing to representation to knowledge discovery and these stages will be broken down and outlined in this section.

Figure 9: Framework of social media text analysis (Aggarwal and Zhai, 2012, p.389)



The core of much of the work in this project will encompass the field of text mining which involves understanding the underlying characteristics of text data. Unlike numerical data which is generally easier to process, text data is by default highly dimensional, unstructured, and can be arranged in many different ways (Fuka and Hanka, 2001). However, the gains in knowledge can result in a richer understanding of the data, especially as more and more emphasis is being placed on getting insights from large text corpora. Different languages also pose an issue when applying pre-processing techniques such as tokenisation and Named Entity Recognition to the data. For example, the EU Parliament which operates in 24 official languages (European Parliament, 2018) could face a difficulty when using text analysis techniques and software. As this project will focus on the UK Parliament, only English will be used and analysed.

As the data to be used in this research is primarily text data in the form of comments posted to social media sites, the first parts of this section will focus on the different methods used to pre-process text data, including ways to handle the high dimensionality. These relate to the pre-processing steps of stop word removal and stemming and the dimension reduction and representation covered in section 3.2.1. Furthermore, supervised and unsupervised machine learning methods including classification, sentiment analysis, and various topic modelling algorithms are explained in sections 3.2.2 and 3.2.3 with regards to how they can be applied to real-world scenarios. Section 3.2.4 covers how I infer the age and education level of users from their text and section 3.3 explains the geo-spatial analysis methods used in the thesis.

3.2.1 Pre-processing and Representation

Although time consuming and tedious, data preparation is a vital step of not just text mining, but all types of data science and analytics techniques (Provost and Fawcett, 2013; Silge and Robinson, 2017). Avoiding this step would result in unreliable data analysis, drawing incorrect and biased conclusions. The first step is usually tokenisation which creates smaller units of words from a document, usually removes any punctuation, and converts all letters to lowercase (McCallum, Wang and Corrada-Emmanuel, 2007).

Stop-words are frequently occurring, words that carry little meaning in a document – for example “the”, “to”, “a”. These are often the first words to be removed to gain a ‘truer’ understanding of the text and meaning (Blei, Ng and Jordan, 2003; Aggarwal and Zhai, 2012). The stop-words usually have a high frequency in a document, and as they carry little meaning they can be misleading when it comes to analyse tasks such as information retrieval and extraction. Some analysis and visualisation techniques such as word clouds can be deceptive should the stop words be included. Word clouds are visual representations of word frequencies with more frequent words appearing larger and less frequent words appearing smaller.

There are many dictionaries/lexicons for stop-words, which can be applied to the data and in effect, filter out the specific words it deems to be stop-words (Blei, Ng and Jordan, 2003; Dreiger, 2013; Silge and Robinson, 2017). With relatively modern and structured 'clean' text in the form of articles, novels, even emails to some extent, these lexicons can be very useful in selecting the correct words. However, for different types of text, for example novels written by 16th century authors, or 21st century use of social media which use a different form of the modern language, the traditional stop word lexicons may prove to be unsuitable. This project will deal heavily with social media data, therefore this is something to take into account, and in this case the popular Snowball lexicon will be used (Snowball, 2020). This lexicon contains a mixture of 571 pronouns (i.e. "I"), verbs (i.e. "were"), auxiliaries (i.e. "should"), compound verbs (i.e. "they've"), negation (i.e. "isn't"), and others.

After removing stop words from the text the next step in the pre-processing process was stemming and lemmatization. Stemming and Lemmatization are pre-processing measures used to break down the words in a document. Stemming then removes any morphological inflections from words. For example, the words *remained, remaining,* and *remains* would have their suffixes removed and be reduced to *remain.* This of course loses some of the semantic qualities of the words, however in many cases the reduction in dimension far outweighs the slight loss in meaning. Where stemming does not take the linguistic unit into account, lemmatization does (Jivani, 2011). The latter will first determine the Part-of-Speech (POS) i.e. verb, noun, etc. before reducing the unit. In this thesis, I use the tm (Feinerer, Hornik and Meyer, 2008) and udpipe (Wijffels, 2019) R packages to handle lemmatization.

N-grams are commonly used to bring to light co-occurrences in the text which could aid analysis and classification. An n-gram has an order (number of entities) and a type (i.e. character, word...). For example, a word-level order-2 (bigram) n-gram will contain a list of the sequential pairs of word co-occurring in a text. Tabulated word-level n-grams can be used to uncover words that most commonly co-occur within a text, and therefore reveal some common themes within a document further than simple word clouds. In my research, I use bigrams to explore the most frequent word pairs and as a first step to uncovering themes in a discussion. I display these bigrams in a network with an arch connecting each word pair and the colour of the arch changing depending on how frequent the word pair is in the corpus. Representing bigrams in this way rather than in a bar chart allows the reader to view how the bigrams overlap and reveals clusters of words which all co-occur together. This can highlight patterns of words which are popular in the discussion and is a preliminary method of showing the main topics within the corpus.

Named Entity Recognition (NER) and Part-of-Speech (POS) tagging are two further methods used in text mining to format the data in such a way that will inform future analysis techniques. NER will highlight the terms in a document which relate to a person or a place. POS tagging will assign a grammatical class to each word in the document – much in the same way a syntactical grammar tree can represent a sentence. This can aid with other analysis techniques such as sentiment analysis in which (in English) most of the sentiment-carrying words are found in adjectives. Tagging text in this way can provide a basis for future feature selection and classes in the context of topic modelling.

Pre-processing social media data also comes with its own unique challenges. Hong and Davison (2010) note that removing URL links to external websites and words starting with the '*@'* symbol (where users had tagged each other) were also necessary before completing any further analysis. Emoticons (i.e. ☺, ☹) and hashtags (#) used in Facebook and Twitter are often removed during analysis (Tian *et al.*, 2017). However, in some cases when dealing with social media data, the existence of emoticons aids in sentiment analysis and the hashtags help to discern the important topics of a message or comment. For example, if an analyst wanted to select any positive tweets they could filter the dataset for all tweets containing the smiley face (☺) emoticon. Furthermore, keeping the mentions in Twitter data can be very useful when analysing a network of interactions between users, so removing these elements of social media data may prove detrimental to future analysis. In section 6.2.1 I use the presence of 'RT' in a tweet (signifying the tweet is a retweet) to create an edges table and complete social network analysis on a dataset of tweets. For this reason, these elements of the comments (mentions, hashtags, and emoticons) will remain for analysis.

One common approach to text mining is using a bag-of-words technique which treats each document in a data set as a vector of word frequencies (Forman, 2007). The order for these words is not taken into account, simply the frequency of the particular word in a particular document. An intuitive way to represent a corpus is through a Document-Term Matrix (DTM) containing the individual documents as rows and unique words (terms) as columns[8]. In the context of this thesis, a document refers to an individual comment posted on a social media platform or online forum. The values in the matrix are the number of times each word appeared in each document. The values can also be changed to a weighted term frequency such as Inverse Document Frequency (IDF) (Blei, Ng and Jordan, 2003; Silge and Robinson, 2017) depending on what is being measured. The Document-Term Matrix can also provide a sparsity percentage which is commonly very high due to the unlikeliness of many documents containing the same words. The dimensions of these matrices are also very large due to the number of possible unique words in a document, especially if stemming or lemmatization has not been performed. Previous research showed that within the first 20,000 words of a novel, almost 4000 of those (20%) were unique and this figure steadily rises proportionally through the trajectory of the book[9]. The size of these matrices can be reduced using the pre-processing techniques above, especially stop-word removal and stemming/lemmatization. Many of the techniques used in the pre-processing stage of text analysis have dimension reduction in mind (Adeva *et al.*, 2014; Jivani, 2011). The highly dimensional and unstructured nature of text makes successfully analysing language very difficult. The theory of generative grammar states there are an infinite number of sentences that will never have been uttered before, and to fully understand language is to still know the meaning of these sentences despite the novelty of them (Haegeman, 2009). This highlights the unpredictability of free text such as those found on social media and is one of the reasons Natural Language Processing (NLP) methods and text mining is so important.

Because of this issue of high dimensional text data, I used dimension reduction to aid text mining. Dimension reduction can be achieved through a process called feature selection (Fuka and Hanka, 2001) which involves selecting specific characteristics of the text which many of the terms fall into. In doing this, a large proportion of the data can be represented by a much smaller subset which can aid in overall accuracy and scalability (Forman, 2007). There are a range of methods varying from simple to more sophisticated using machine learning models. A very simple method of feature selection is simply setting a threshold value, outside of which words will not be used. Taking the 50 most frequent words in an article for example would already provide much more useful insights than the whole text (Vakulenko, Nixon and Lupu, 2017). Visualising these in a word cloud also makes it quick and easy to uncover some topics.

---

[8] Note, this can also be shown as a Term-Document Matrix with the contents of the rows and columns reversed.
[9] Prior research done by author

A word cloud displaying simple term frequency is useful, however when dealing with many documents with different topics and different authors, often a weighted term frequency is more appropriate. The most common used words in a text may not be the ones which make the text most identifiable among a corpus. For example, I have seen that stop-words are often the most frequently occurring terms in a document, however searching for these would not be helpful when trying to differentiate between different Facebook comments. Rather, searching for nouns or verbs more specific to a particular comment would lead to more relevant and useful results. TF.IDF is a measure which weights term frequency (tf) with the inverse document frequency (idf). The higher the IDF score, the less frequently the word appears in the document giving stop-words very low scores and the more specific words higher scores (Blei, Ng and Jordan, 2003; Silge and Robinson, 2017; Forman, 2007).

Although at first seeming counter-intuitive, when weighted against each other, these two measures of term frequency and inverse document frequency form some of the underlying theory and algorithms behind information retrieval applications such as web search. These applications aim to return the most likely documents for a specific word in a corpus rather than a document which mentions that word the most frequently. Formulas below outline the process of calculating this measure.

$$tf = \frac{t_d}{|Vocabulary|}$$
$$df = |d_t|$$
$$idf = \log\left(\frac{|d|}{df}\right)$$
$$tf.idf = tf \times idf,$$

Where t_d is the number of times a term t appears in document d, $|d_t|$ is the total number of documents with a specific term t, $|d|$ is the total number of documents in the corpus, and $|Vocabulary|$ is the total number of terms in the document.

Subsets derived from Named Entity Recognition and POS tagging can also be used as features to successfully classify documents. The Document-Term Matrix (DTM) representation had the disadvantage that it is usually very sparse and of a large dimension. Throughout the course of this thesis I will be using various methods of feature selection to subset my data and provide more interpretable results. For example, I use the tf.idf measure as a sorting variable when displaying the words most associated with a particular topic in a topic model. This approach is helpful when analysing topic models of digital discussions because it reveals the most unique words in a topic relative to the words in other topics. For this reason, when building 'TheGist' application for text mining, I include an option to display topic models using the tf.idf measure, which gives precedence to words most unique to that topic. This is explained further in section 3.2.3. Another form of feature selection used in this research involves POS tags. Research suggests different age groups use different grammatical features in their writing, and this is used to infer the age and educational background of participants to online discussions. I use POS tags to first identify and then calculate the prevalence of nouns, adverbs, pronouns, and interjections in a comment, so I can estimate the age of the participant. Further details are explained in section 3.2.4. A final method of feature selection is through word associations. The findAssocs() function in R takes a DTM, a word, and a value for the lower correlation. This returns a list of words in the DTM which are correlated with the word specified in the function call. Visualising this in a word cloud where the size of a word is relative to its correlation with the original word allows me to easily view which words are often used together in a different way to a bigram network.

3.2.2 Knowledge Discovery - Sentiment and Emotion Analysis

Once text data has been sufficiently pre-processed, the possibilities for analysis are endless. One common technique used is that of sentiment analysis. It aims to uncover the underlying sentiments behind text, much like humans naturally do when reading a document (Silge and Robinson, 2017, p.13). For example, understanding whether the text is generally positive or negative is very useful when analysing film or hotel reviews and in any case where the opinion of the authors of the text is valuable information. Therefore, in cases such as parliamentary public engagement activities where understanding the public's opinion on certain matters is vital to measuring the success of the activity, sentiment analysis provides a good starting point.

Sentiment analysis is already frequently used in social media contexts when understanding situational awareness of tweets or even understanding the political leanings of users (Sen, Rudra and Ghosh, 2015; Verma *et al.*, 2011; Nielsen, 2011b; Pu, 2017; Maynard *et al.*, 2017; Pennacchiotti and Popescu, 2011). The majority of sentiment analysis exercises follow a supervised learning model, where pre-annotated data is subset and trained. This enables the model to learn based on pre-annotated training data and then be tested on another pre-annotated subset of the data. It is vital that this second test set is completely distinct from the training set, as any overlap will unfairly skew results and misleadingly raise the accuracy of the model. This leads to the overfitting of the model (Provost and Fawcett, 2013). The reason for the pre-annotation of the subsets with the correct sentiment classification is so they can be accurately checked against the model's predictions. An accuracy score of the model can then be calculated based on the number of correct predictions it makes on the test data with measures such as precision and accuracy. Once refined, the model can be applied to unseen, un-annotated data. While sentiment analysis using machine learning is a useful and established method, I chose not to train a new model in my research. This is because I found using a sentiment dictionary or lexicon (created from a read-trained model) to provide accurate results without the increased computational cost of training many machine learning models for different datasets. Furthermore, training a supervised learning model can be computationally expensive, as they require a portion of the data to be manually pre-annotated with the correct classification. With a large dataset, this could prove time consuming in the beginning as a manual pre-annotation of the training and testing data is required, therefore a semi-supervised sentiment lexicon approach may prove more time efficient. Using a lexicon enables the scaling of the data, meaning when there is a large volume of input it would still be quick and easy to uncover the important sentiments from the data which can in turn be passed onto policy-makers. Furthermore, due to the nature of text data, many sentiment analysis exercises do not capture irony or sarcasm very well. There is other research to automatically detect irony and sarcasm for example in Maynard and Greenwood (2014) and Wiegand *et al.* (2010) but this remains outside of the scope of this thesis. With any data including social media data the opinions people express are generally very specific to what they are discussing, and may therefore struggle to be generalised to a wider/broader topic (Maynard *et al.*, 2017).

There are many sentiment lexicons available for sentiment analysis, the choice of which can affect the final results. Previous research found that comparing six different sentiment lexicons to Twitter data yielded different results, and lexicons made up of larger word lists did not actually yield better accuracy results (Nielsen, 2011b). Through using a sentiment lexicon, it's also possible to determine which words contribute most to the positive or negative sentiments in text, something that may prove very useful if trying to understand which topics are garnering particular feelings during an online consultation or participation activity (Silge and Robinson, 2017). I use three different sentiment lexicons, Bing, NRC, and Afinn, within this thesis because they each provide a different interpretation of the data and are useful in different contexts.

The Bing sentiment lexicon is binary positive, negative and assigns a category to each word in a document also listed in the lexicon (Liu, 2018). This lexicon has around 6800 words and is very useful to get a preliminary overview of the main sentiments of a discussion. It is also useful in terms of visualisation to create a word cloud with the positive and negative sentiments differentiated by colour. The National Research Council of Canada (NRC) EmoLex lexicon is also categorical using two sentiments of positive and negative, and six specific emotions; anticipation, anger, joy, sadness, trust, and fear (Mohammad and Turney, 2013). This lexicon has over 14,000 words and adds a richer interpretation of the data because I can identify precisely which words participants use to express specific emotions. Using NRC the discussions can also be subset depending on the emotions being expressed, for example one could examine 'fear' and 'sadness' comments of a discussion. Finally, the Afinn lexicon uses a 10-point scale from -5 (very negative) to +5 (very positive) to categorise the words in a document (Nielsen, 2011b). This lexicon has over 3000 words and provides a different interpretation to Bing and NRC. The categories in this lexicon capture the intensity of sentiment rather than categorising them based on the type of sentiment expressed as the other lexicons. For example, curse words usually appear in the most negative category (-5) so Afinn provides a good way to separate comments with a lot of cursing which can suggest participants with little to add to the conversation or even possible trolls.

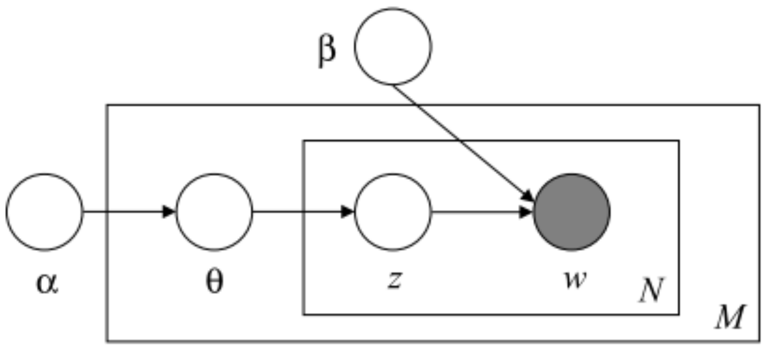
In all three cases, the sentiment of a comment is calculated based on the sum of all sentiments in a document proportional to the total comment length. For example, if a sentence has 3 negative words and 5 positive words identified by the Bing lexicon, its final category will be positive. In the case of the NRC lexicon, a comment can be classified as both sadness and anger for example. Many tools are available for applying sentiment analysis to a dataset, the tidy text mining package in R (Silge and Robinson, 2016) provides access to all three sentiment lexicons and I use ggplot2 (Wickham, Chang and Wickham, 2016) for the visualisations.

3.2.3 Knowledge Discovery - Topic Modelling

My aim is to understand what people participating in the parliamentary online engagement activities are saying with respect to a given topic in relation to the spectrum of engagement, therefore, uncovering the topics from data. Hong and Davison (2010) describe topic models as "…powerful tools to identify latent text patterns in the content." (p.80). More broadly, they are "used to discover a set of "topics" from a large collection of documents" (Wang and Blei, 2011, p.450). This can of course prove very useful in cases of large-scale online public engagement activities where an overview of important themes is useful without manually reading through each comment. This section will provide an overview of topic models that will be used to analyse the data resulting from UK Parliament online engagement activities.

The probabilistic algorithm called Latent Dirichlet Allocation (LDA) topic modelling approach is commonly used in social media settings and classification. In LDA, each document is represented as a probability distribution over a number of topics, and each topic is a probability distribution over a number of words (Hong and Davison, 2010; Blei, Ng and Jordan, 2003) The model is explained visually in Figure 10 (Blei, Ng and Jordan, 2003, p.997). More clearly, the shaded set *w* represents the only observable entity of the model, words. The topic assigned to the word by the model is represented by *z*, both of which are repeated for *N* number of words in a document. This is then repeated for the corpus of documents *M* for which a mixture of topics θ is assigned. The external or corpus-level parameters α and β control the per-document topic distribution and per-topic word distribution respectively. A high value for α suggests most documents contain a mixture of different topics and the documents appear more similar to each

Figure 10: Representation of LDA Model using plate notation



other. A low α value suggests the documents are represented by only a few topics. In the same vein, a high value for β suggests topics are likely to have a mixture of many different words and therefore topics appear more similar to each other. A low β value suggests only a few words make up a topic (Knispelis, 2016; Blei, 2017).

I implemented the topic model using the LDA function of the topicmodels package in R (Hornik and Grün, 2011). This requires the data to be in a bag of words format which is represented as a document-term matrix (DTM) which contains words in the columns and documents in the rows. It requires the analyst to set the number of topics *k* in advance. I chose VEM which uses the maximum likelihood method to estimate the parameters for α and defaults to 50/k. The function returns a fitted LDA topic model containing a beta matrix of the per-topic word distributions with dimensions *k* x *w* and a gamma matrix of the per-document topic distributions with dimensions *M* x *k* (Hornik and Grün, 2011).

As it requires no training or pre-annotation of the data for the model, LDA is an unsupervised learning method and can uncover several topics in the same document. This is more representative and reflective of real-world data, and noted by several researchers as an advantage which sets the model apart from simple clustering (Wang and Blei, 2011, p.45; Blei, Ng and Jordan, 2003, p.997). To decide on an appropriate value for *k* in the LDA topic model, the documents are run through the "LDATuning" package which evaluates different LDA models each with a different value of k (i.e. from 2 to 20) based on a set of 3 measures (Nikita, 2016): the maximum extremes of the Griffiths and Steyvers (2004) index, and the minimum extremes of the Cao *et al.* (2009) and Arun *et al.* (2010) indices. Validation of topic models is a well-documented issue in the literature due to the contrast between optimum topic categorisations as defined by various measures (i.e.coherence, perplexity) and the interpretability of topic distributions by the human (Maier *et al.*, 2018; Nikita, 2016; Röder, Both and Hinneburg, 2015). For example, Vidgen and Yasseri (2020) validate their LDA topic model using word intrusion which requires subjects to identify the odd-one-out word within a topic (Chang *et al.*, 2009). They ask three students to decide which one word out of six has the lowest probability of being in a particular topic. In this way, students identified the anomaly word across a set of words in a topic as a method of validating the manual topic labels chosen by the authors. I did not use this method as I was working alone on this project, and found I was able to sufficiently validate the model qualitatively. Additionally to using statistical measures as defined by Griffiths and Steyvers (2004), Cao *et al.* (2009) and Arun *et al.* (2010), I used a qualitative approach, by judging the validity of the models based on their interpretability. This was particularly critical for my collaborative partners who valued their ability to qualitatively understand the topics over the statistical validation measures. As a result, I use a combination of statistical measures and qualitative analysis to select the most suitable number of topics for a given model. However, discrepancies in validation of topic models is a limitation of this particular method.

Contextualising this topic model within this project, the documents represent different comments made by participants of a digital debate for example. The model would then allocate a topic to each word used and then assign a topic probability distribution to each comment. This would be analysed by whichever measure of distance chosen and boundaries drawn to separate the different comments based on their topics. The topics and words appearing in the topics are displayed in a graphical or tabulated format could then be easily analysed to understand the main topics in the discussion without the need to read each comment separately.

In my research I use several methods of visualisation of topic models. The first is showing the distribution of the number of documents in each topic. This is useful for analysis as I can see whether the majority of documents are categorised into one single topic, or more spread out between a range of topics. I visualise this through bar charts. The second form of visualisation is for the beta matrix of the per-topic word distributions. By using this matrix I can create separate bar charts of the words with the highest beta scores in each topic. This provides a good overview of the primary words associated with each topic and allows me to understand what each topic is about. The final form of visualisation of topic models is through the LDAvis package also in R (Sievert and Shirley, 2014). This uses the beta and gamma matrices, a list of terms in the topic model, the length of each document, and an approximation of the distance between topics. I used the Principal Component Analysis (PCA), which involves explaining the variability of a dataset with fewer dimensions. This smaller dimension dataset enables me to view the variation among the variables along a single dimension. This method uses eigenvalues and both covariance and correlation matrices to derive the proportions of variables that affect the total variability. These are referred to as the components, and the loadings of the components determine which of the original variables contribute the most to the variation. These loadings are the eigenvectors. LDAvis outputs an interactive visualisation with an intertopic distance map on the left and a bar chart on the right. The distance map plots the topics as nodes on a graph with the nodes separated by the first two principal components. The size of the nodes is relative to the number of documents in each and nodes positioned close to each other are more similar. On the right of the visualisation lies a bar chart with the 30 most frequent words in the whole corpus. Selecting one topic from the left changes the bar chart to show the 30 most frequent words in that particular topic. This interactivity is a useful way to explore the topics in a discussion and to see how the word usage changes depending on the topic chosen. There is also an option to alter the relevance metric, λ, in the visualisation which changes the order of words in the bar chart. This λ ranges from 0-1, where higher values sort the words according to their probability of belonging to a particular topic, and lower values sort the words according to a ratio of topic probability to probability of occurring in the whole corpus. Sievert and Shirley (2014) and Taddy (2012) note that this measure "generally decreases the ranking of globally frequent terms" (p.65) which bears resemblance to the tf.idf measure introduced in section 3.2.1. The LDAvis visualisation is embedded into TheGist application and I recommend users to set λ to 0.6 to still have a compromise between viewing the terms with the highest topic probability and ones that are also unique and most relevant to the selected topic.

Document length is also something to bear in mind when working with topic modelling. The majority of applications require the number of words per document N to be rather large, however in a social media and microblog setting, the documents (represented by comments or tweets) have a small value for N. Therefore, training the LDA topic model on short comments may decrease its ability to assign accurate probability distributions for topics. This is due to LDA's model architecture. With a limited number of words per documents it is difficult for the topics to be identified as they may only have a small number of words associated with them, out of an already small set of words. For this reason, several methods of word and bigram frequencies, and LDA topic models are used to uncover themes from online discussions.

3.2.4 Knowledge Discovery - Estimating user socio-demographics from text

For the purpose of this research, understanding what kinds of people are participating in the parliamentary online engagement activities is as important as understanding what is said by them. The introduction of social media created a new avenue for many organizations, including political bodies, like parliaments and governments, to inform the public of their news. But these media also allow institutions to harness opinions from a wide range of people in the form of product reviews or general comments. However, along with other difficulties created by the use of large volumes of unstructured textual data, it is difficult to know who is involved in these online discussions. This is a problem that has affected the UK Parliament which uses comments posted online to better understand public opinion on certain topics (Liaison Committee, 2015). Less than half of the population think that they can influence decisions by engaging with parliamentarians either online or offline (Hansard Society, 2019). There is a difference in the types of people who engage. For example, it is generally understood that the 'usual suspects' are those who are middle class and well educated, and tend to engage most with Parliament, however studies have shown that when the channel of engagement is moved online 18-34 year olds may be the more engaged (Hansard Society, 2018). Other institutions have focused on engaging sectors of society who fall outside of the 'usual suspects' category including young people (Lee and Young, 2013; OurSpace, 2017; Inter-Parliamentary Union, 2016). This is often achieved by directly targeting young people with or by engaging with communities where certain less engaged segments of societies are already involved.

It is important to know who is engaging to understand who is excluded, so they can be better targeted. One way to identify the socio-demographic background of a person online could be to analyse their writing style. There are many factors which can influence differences in one person's writing style including the subject matter or the audience (Brandt et al., 2020; Sloan et al., 2015). One person writing a full-length academic paper may have a very different writing style compared to when they are writing a tweet or commenting on a Facebook post. Knowing the education level and income status of this individual from other data sources, would allow me to correlate the presence or absence of syntactical features in their writing with their socio-demographic attributes. But obtaining all this information for people online in a parliamentary setting this task is complicated because I do not have access to this supplementary information from participant of digital discussions on Facebook or Twitter.

It is therefore important to attempt to find ways to infer the socio-demographic background of online participants as accurately as possible, using only the digital traces, that these participants leave behind as they engage. Syntactical features are used to develop an understanding of users who contribute to online discussions and specifically to better understand their socio-demographic background. Prior research has identified the use of certain syntactical features in writing as good indicators of age, income level, and education level of users (Flekova, Preoţiuc-Pietro and Ungar, 2016).

Assessing the education level of participants can be achieved through the Flesch Readability score (equation 1) and the Flesch-Kincaid Grade Level (equation 2). These have been used to measure how easy or difficult a piece of text is to read and approximately what level of education is required to understand the text (Flesch, 1948). The Flesch-Kincaid Grade Level had been introduced to transform the Reading Ease scores to the me Grade system. This provides further context and aids interpretability for the Reading Ease scores and can be used to determine the average education level of the participants online (Kincaid *et al.*, 1975). Both measures take the average sentence length and number of syllables in the word into account to calculate a range of reading ease scores (from 1-100) or US grade levels (1-12). In order to make these calculations more relevant to the UK context of this research, I have amended the Flesch-Kincaid grade Level equation (2) to reflect the UK school level system (1-13) (equation 3).

$$206.835 - 1.015\left(\frac{total\ words}{total\ sentences}\right) - 84.6\left(\frac{total\ syllables}{total\ words}\right) \tag{1}$$

$$0.39\left(\frac{total\ words}{total\ sentences}\right) + 11.8\left(\frac{total\ syllables}{total\ words}\right) - 15.59 \tag{2}$$

$$\left(0.39\left(\frac{total\ words}{total\ sentences}\right) + 11.8\left(\frac{total\ syllables}{total\ words}\right) - 15.59\right) + 1 \tag{3}$$

Another way to estimate demographic information, in particular age, from users comes from analysing the underlying linguistic attributes that make up text, for example, the prevalence of nouns, adjectives and pronouns. These attributes can be extracted automatically from text through part-of-speech (POS) tagging, which is a common pre-processing stage of natural language processing as described in earlier sections. Previous work using these POS tags identified the presence of 3rd person pronouns, determiners, adjectives and conjunctions as indicators of older age, and the presence of nouns, interjections, adverbs and 1st person pronouns as indicators of younger age (Nguyen, Smith and Rosé, 2011; Brandt et al., 2020). These attributes are calculated in relation to total length of each comment (Flekova, Preoţiuc-Pietro and Ungar, 2016). In this thesis, I create an age indicator which has a value between -1 and +1 for each comment.

Comments with an age value closer to -1 have a higher prevalence of POS tags suggesting the author is an older age and comments with an age value loser to +1 have a higher prevalence of POS tags for younger age.

Research has identified correlations between the presence or absence of certain linguistic attributes and the age and income of the person who wrote the text (Flekova et al., 2016). This research has also found the Flesch Reading Ease score to correlate with the age and income of the authors of a piece of text more than any other estimation. Under these assumptions, it is possible to infer some socio-demographic characteristics from the comments posted by people engaging online using social media. Specifically, through a combination of these reading scores and age indicators, I can create demographic profiles of each comment in a discussion. Using the results from the readability scores and part-of-speech analysis (age indicator), users are classified in various groups (e.g. low education, high education) using clustering analysis. I apply k-means clustering (Kaufman and Rousseeuw 2009) to the comment data incorporating age indicators (identified through POS tags) in combination with the UK grade levels (adapted from Flesch-Kincaid Grade Levels). K-means clustering "is an iterative method which minimises within-class sum of squares for a given number of clusters" (Charrad et al., 2012, p. 18). It is an unsupervised method which requires the analyst to select the number of clusters k to compute. The clustering was run through several optimization measures to evaluate the most suitable value of k. The K-means clustering was completed using the cluster package in R (Maechler, 2019) and the optimization of cluster numbers using the nbclust package (Charrad et al., 2012). This package computes the Elbow, Silhouette (Kaufman and Rousseeuw, 2009) and Gap statistic (Tibshirani, Walther and Hastie, 2001), each method providing a plot of a range of number of clusters with an indication of which value of k is optimal. As with obtaining the optimal topics for LDA topic modelling, I used these optimisation measures for k-means clustering and the results are explained in detail in section 6.2.3. After identifying the clusters, I apply again sentiment analysis and LDA topic modelling of the posts of users in the separate clusters in order to better understand whether these various socio-demographic groups vary in the way they contribute to political discussions initiated by the UK Parliament.

3.3 Geo-spatial analysis

In section 3.1 I explained that the data from certain social media platforms, specifically Twitter, includes user location information which can be plotted onto maps. This is a helpful form of analysis as I can see where in the country participants are based and whether the digital discussions are reaching a wide range of the population. In section 5.2 I plot the locations of Twitter users (based on the information they provide on their Twitter accounts) for followers of the 48 of the UK Parliaments Twitter accounts. I created a new csv file using the locations by city (I only used entries which included the city name, and omitted the country name from those users who included it), the number of combined Twitter users in that city, and their number of combined Twitter followers and Twitter friends (people who follow them). However, there are also pitfalls in using social media data for geo-spatial analysis. This is especially the case in terms of location tags for tweets where accurate geographic location information is not available in the free Twitter sample. Users on Twitter have the option to add their own location in their profile description in free text form with no data consistency. This becomes very difficult to map as it may not be accurate e.g "The Red Rose Empire", several cities e.g. "Mexico City/Toronto/London", different categorisations e.g "Swansea" vs "Germany", or simply not given. I used the data visualisation software Tableau (Tableau.com, 2020) which includes a mapping feature including the coordinates for cities around the world. I linked these coordinates to the cities in my dataset to create the maps where the size of each point on the map is relative to the number of Twitter users in that city. The inclusion of the follower and friend data into the dataset allowed me to also alter the colour of the points depending on how many combined friends or followers the users in a specific location have. This allows me to assess whether the locations with high numbers of Twitter users also themselves have many followers and friends. This can highlight areas of the country where there may be a relatively small number of Twitter users in comparison to others, but where those users have a high number of friends. These location with a high user to friend ratio could be areas to target when aiming to spread news of an inquiry launch as anything retweeted by those users would be viewed by a large number of their friends.

In section 6.6.1.1 I also use geo-spatial analysis with the Discourse dataset. Purpose-built engagement platforms such as Discourse provide the IP (internet service provider) addresses of each registered user, from which latitude and longitude coordinates can be derived and plotted. With this information, I follow the same steps as above to plot the locations on a map using the Tableau software. These differences in data veracity alter the types of analysis that can be done on both Discourse and Twitter platforms, and as a result what insights can be gained. Additionally, as an aim of parliamentary engagement is to reach and hear from people in all areas of the country (Liaison Committee, 2015), using tools which enable this type of analysis serves as a great advantage to the evaluation of the activity.

3.4 Social network analysis

The study of networks is useful in many disciplines including many of the natural and physical sciences: “Social network analysis (SNA) is the study of mathematical models for interactions among people, organizations and groups” (McCallum, Wang and Corrada-Emmanuel, 2007). For the purposes of this research, understanding the links between people who interact with online engagement activities will help to provide a richer understanding of the data and in turn the evaluation of digital engagement. Furthermore, the differences between social networks resulting from different online engagement activities can be analysed to help inform policy-makers on the public’s interaction with each other during digital discussions and how communities of users form based on the subjects they are discussing. Social network analysis can be applied to study relationships between various entities, ranging from individuals, over businesses, to countries. The focus here will be on interactions between social network accounts, which usually represent individuals, but can also include interest groups or organisations.

Social network analysis can have different terminology in the literature to describe different elements of the graph, however for the purposes of clarity only the following will be used. A *node* represents a single entity, as can also be referred to as an actor. In many visual representations of graphs, a node’s size is a reflection of their ‘importance’ in the network, and a larger node represents one that has many connections to other nodes and therefore a high eigenvector centrality (Asher, Leston-Bandeira and Spaiser, 2017). Nodes are connected by *edges*. These represent the presence of relationships or connections between nodes. They can be directional signifying a relationship may not be mutual (for example one Twitter user may follow another but may not be followed back). Depending on the context of the graph, the direction of the relationship might not be relevant, for example if one country trades with another it is automatically reciprocated. A *community* is a concentration of nodes which are strongly connected to each other. Identifying these can be very useful in exploring the data to find which Twitter users are most connected to each other along a specific measure, for example, users that retweet each other, or people who are Facebook friends. Just as text data can be vectorised in a document-term matrix, network data i.e. data that captures entities and edges between entities can also be used to construct an adjacency matrix. This is a square matrix with each node appearing in the rows and columns, and the values representing the edges between them. In a binary setting, the value 0 represents the lack of a relation or interaction between two nodes, while a 1 represents a connection between two nodes. The diagonals will have 0 as values, as a node cannot be associated with itself (Dean, 2018).

Measures such as modularity and centrality are common within the literature of social network analysis and are used to measure distances between nodes and the importance of certain nodes compared to others (Asher, Leston-Bandeira and Spaiser, 2017; Dean, 2018; Kolaczyk, 2009). An in-depth explanation of modularity will be given in the following section in relation to the Louvain method of community detection. Centrality of nodes can be measured in various ways however I will be focussing on betweenness centrality. *Betweenness Centrality* is concerned with where nodes lie on a path in relation to other nodes, and how many paths they intersect - a node which intersects many paths will have a high degree of betweenness (Kolaczyk, 2009; Asher, Leston-Bandeira and Spaiser, 2017). Used in tandem and with the textual content of the data in the form of comments, this measure of centrality helps to interpret the analysis of the network graphs.

Smith *et al.* (2014)'s Pew Research Centre report outlines several characteristics of networks that they have found through analysis of Twitter data (Figure 11 (Smith *et al.*, 2014, p.8)). These networks fall into 6 categories, categorised by the nature of their entities and

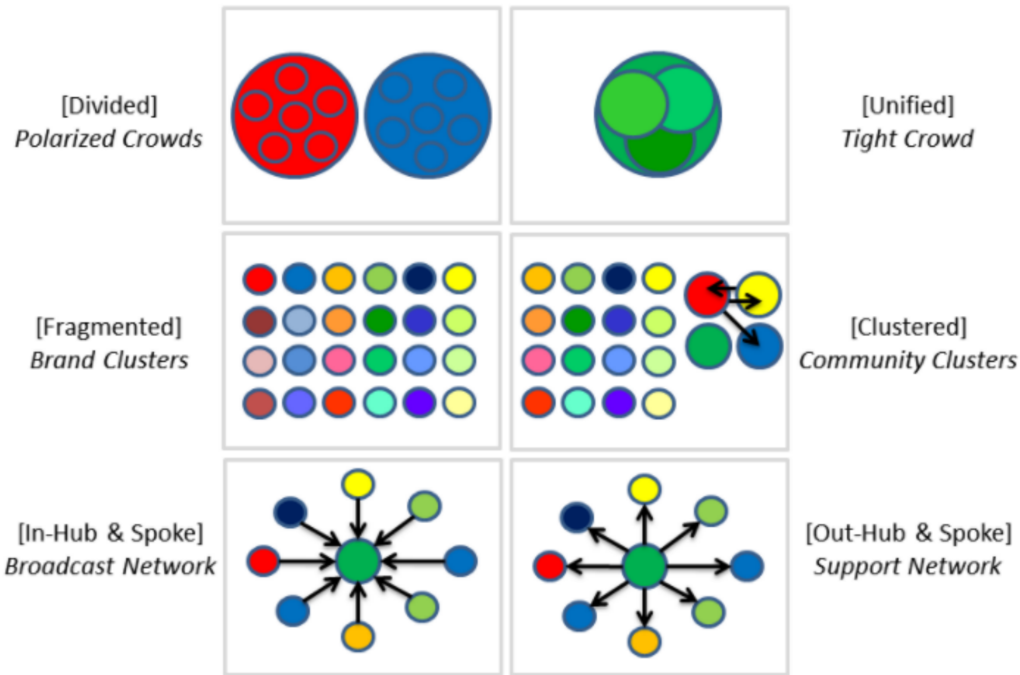


Figure 11: Diagrams of 6 types of network (Smith *et al.*, 2014)

relationships; Polarised vs. Tight Crowd; Brand vs. Community Clusters; and Broadcast vs. Support Network.

Smith *et al.* (2014) argue the different types of network suggest different characteristics about a certain network and the people participating in it. For example, the difference between a heated political discussion in which neither side really agrees or interacts with the other, and one that encourages cross-communication can be described by a *polarised* or *tight crowd* respectively. The *brand* vs *community* cluster distinction represents a situation where users in a discussion form their own small communities independent of each other with a large number of unconnected participants, compared with *community clusters* which have fewer unconnected participants showing users are interacting with each other albeit in smaller numbers that the polarised or tight crowd networks.

As network graphs can be directed, two further distinctions of a *broadcast vs. support network* can be made. Communication flowing into one central hub in Twitter can be explained by many users retweeting one influential user, for example when an international news story breaks. Interestingly Smith *et al.* (2014) found that these users do not have many interactions in common, other than the fact that they retweet the same account. A similar pattern was found by Asher, Leston-Bandeira and Spaiser (2017) in their analysis of e-petition debates on Twitter. Their findings suggest that during a parliamentary debate about the banning grouse shooting e-petition on which Twitter users were able to comment, many users were retweeting one more central and influential account – that of Natalie Bennett who was the former leader of the Green Party.

Conversely, when one central hub is interacting with many different users a *support network* is formed. Smith *et al.* (2014) characterise this with an example of a company dealing with customer complaints, the central company account replies to many different customers who have tweeted them. This could easily be imagined with airlines who have delayed flights and need to respond to frustrated customers, or an electronics company which may need to recall certain products and respond to customer concerns.

These categories are by no means completely representative, however they do manage to capture the features of many types of networks. Smith *et al.* (2014) also explain “many social media topics exhibit a hybrid network structure that combines elements of the six network types described here.” (p.5). Therefore, allocating one type to the data may not be possible, this is especially the case with the Broadcast and Support Network types as their main distinction is in the direction of their edges.

3.4.1 Algorithms for community detection

Having data about user interactions is very useful especially when applied to community detection algorithms. This allows me to explore how different users group based on how they have interacted with each other. It reveals groups of users and identifies any isolated users who are separated from the discussion. Many different algorithms are used for the purposes of community detection in social networks (Bedi and Sharma, 2016; Papadopoulos *et al.*, 2012; Zalmout and Ghanem, 2013). The choice of which to use can depend on the type of network (i.e. directed or undirected), the method of learning (supervised or unsupervised) and the nature of the source data (i.e. overlapping communities, topics, users). The computational cost is also something to take into account (Papadopoulos *et al.*, 2012; Bedi and Sharma, 2016). Due to the nature of social media, unsupervised community detection/graph clustering algorithms generally provide the best results as it can be difficult to estimate the number of clusters/communities in a given network beforehand. Attempting this may well cause the user to miss important clusters (Papadopoulos *et al.*, 2012).

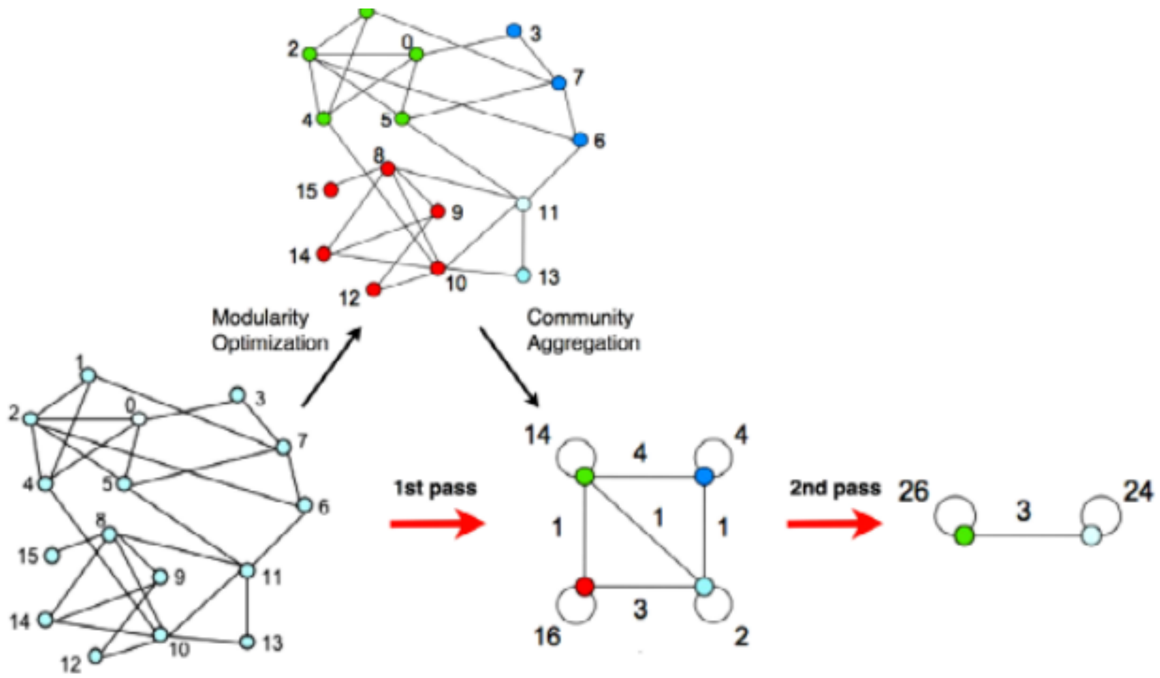
The Louvain method is one algorithm commonly used in social network analysis to identify the clusters of nodes characterised by the six types described in the Pew Research Centre's report (Figure 11). It uses a measure of *modularity* which "assigns nodes to clusters when their connections are above a random chance of connection" (Cihon *et al.*, 2016, p.8). This allows the clusters to form based on the strength of the connections. The equation for modularity is below:

$$Q = \frac{1}{2m}\sum_{i,j} \left[A_{ij} - \frac{k_i k_j}{2m}\right] \delta(c_i c_j) \tag{4}$$

Where A_{ij} is the weight of the edge between i and j, $m = \frac{1}{2}\Sigma A_{ij}$ number of edges, and $k_i = \Sigma_j A_{ij}$ the sum of all weights of edges attached to node i (Blondel *et al.*, 2008). The value Q ranges from -1 to +1 with values closer to 1 suggesting a high modularity and a good quality partition.

The Louvain method then aims to maximise this modularity as much as possible to create comprehensive clusters in the data. It is an iterative process (where the output of one process becomes the input to the next and which is repeated) visualised in Figure 12 (Blondel *et al.*, 2008, p.5).

Figure 12: Iterative Process of Louvain Method for community detection (Blondel *et al.*, 2008, p.5)



This algorithm works in two main phases which are repeated in an iterative process until the modularity is fully maximised. The first phase *Modularity Optimization* assumes each node in a network is a separate community. The modularity of a node *i* is measured when removing it from its own community and placing it in its neighbour's *j* community. If the modularity is higher in this instance, *i* will move to the new community, and the process is repeated until there is no further gain in modularity. The second phase *Community Aggregation* involves building a new network based on the communities created from phase one. Each community forms a new node (represented by the respective colours in Figure 12) (Blondel *et al.*, 2008; Kido, Igawa and Barbon Jr, 2016). The edge weights are calculated by the sum of all edges between communities. In Figure 12, each of the local edges have unit weight of 1. The green and blue communities have edges between nodes 1 and 7, 0 and 3, 5 and 7, and 2 and 6, therefore the weight of the edge between the green and blue node is 4. During this step, the number of nodes in communities increases and therefore the overall number of communities decreases.

This algorithm has been used in various research mainly focussing on social media data in a political setting. Cihon *et al.* (2016) for instance, use a two-mode (bipartite explained in section 3.4.2) undirected network to capture both the behaviour of Twitter users and petitions, and the relationships between them. They use the Louvain method to conclude that due to the lack of strong clusters uncovered by the algorithm, Twitter users tweet about a range of different topics of petitions. Some researchers have found using relative modularity of a network to be more accurate when making comparisons. This relative modularity takes into account the different densities of the networks to provide a more accurate score of modularity. This is found through calculating Q_{max} for each community in the network as defined by Louvain modularity, $Q_{max} = \sum_{K=1}^{K}[L_k/L - (L_k/L)^2]$ where L_k is the number of edges between nodes in a community k and L is the total number of edges in the network and then calculating the relative modularity (aka fragmentation) as defined by $Q_{rel} = Q/Q_{max}$ (Sah *et al.*, 2017; Wachs *et al.*, 2019). I use this calculation in section 5.3.4 to compare select committee evidence networks.

3.4.2 Analysing bipartite and dynamic networks

While the majority of the social network analysis completed in this thesis involves static network data capturing a particular snapshot in time, I also analyse a dynamic network which captures user interactions with posts made to the House of Commons Facebook account over a 20-month period. The analysis of a dynamic graph requires extra processing of the data and the use of algorithms to handle the data. I began by assigning different topics (though LDA topic modelling) to each of the 600 posts made by the House of Commons Facebook account followed by creating a projection of the bipartite graphs. Bipartite networks are a class of network containing different types of nodes (Cihon *et al.*, 2016). In my context, these nodes represent users and posts. These projections resulted in a user projection where users have a connection if they have interacted with the same post, and a post projection where posts have a connection if they have been interacted with by the same user. The ID codes for posts and users were also renamed and shortened to facilitate later interpretation of result and anonymisation. This step is sufficient for standard static networks, but for a dynamic network separate projections need to be computed for each of the 20 months in the dataset between May 2016 and December 2017. Node labels for each month were also created whereby each user node had a link to one single topic defined through LDA topic modelling. This was used within the EVA algorithm to assign a primary topic to each of the communities of users found for each month. The algorithm is an extension of the Louvain algorithm and uses a measure, alpha, which takes into account the modularity using Louvain and the purity (equation 5) of the community based on the shared attribute (in my context topic number) (Citraro and Rossetti, 2019).

$$P = \frac{1}{|C|}\sum_{c \in C} Pc \qquad (5)$$

Where P ranges from 0-1 and is maximised when all the nodes belonging to the same community C share a same attribute profile".

$$Z = \alpha P + (1 - \alpha)Q \quad (6)$$

An alpha of 0 signifies communities created solely based on Louvain modularity and an alpha of 1 signifies communities created based on shared topic numbers. For each month, alpha values of 0, 0.7, 0.8, 0.9, and 1 were extracted from the algorithm, each containing a different number of communities, however only one value of alpha could be used for each time stamp. To decide on the best value of alpha for each month, I maximise the modularity and purity of the communities, and examine the topic numbers assigned to each community. Care was also taken not to select alphas which grouped only a few user nodes into a community. This process was also repeated for different time periods as some months had very few nodes and very small communities. In order to see how the community membership changes over time, a measure of jaccard similarity (equation 7) was used to see which communities in which months had the highest number of shared users (identified by unique codes). This would allow me to examine if users interested in one topic at time T, were also interested in the same topic at time $T+1$.

$$Jaccard(X, Y) = \frac{|X \cap Y|}{|X \cup Y|} \quad (7)$$

Here X is time T and Y is time $T+1$.

The final task was to take these results and create a visualisation that accurately captured the patterns between time periods and topics. For this, I adapted a Sankey Diagram which required some further processing of the data. Specifically, in order to create the diagram, I omitted those users who had only interacted in one time period which reduced the dataset by 87%. The results of this analysis are shown in section 5.3.2.

3.4.3 Measuring Homophily

Homophily "can be defined as the tendency of people to associate and bond with other people who share similarities with them" (Zalmout and Ghanem, 2013, p.84). Network graphs can accurately visualise the network of people participating in a discussion and as such bring to light information about participant behaviour. For example, homophily can be identified by certain patterns of nodes and edges in a graph, e.g. a *Polarized Crowd* type in Figure 11. This network type shows a set of nodes tightly connected into a community which could be based on their similarity to one another. The edges in this community could represent shared likes or retweets of posts concerning a particular topic of interest. A network where users are interacting more with those in their own cluster than in another cluster is good evidence of homophily. Looking at those who follow the users can also shed some light, for example if a user tends to follow users, who are more similar to them, then they are displaying homophily. Both of these patterns have been found using network graphs.

In a study of Twitter data using 40 million nodes, Colleoni, Rozza and Arvidsson (2014) found for instance that the me Democratic party had more evidence of homophily than the Republican party, despite the latter's official account being more popular. This conclusion was drawn from the characteristics of the graph they derived from the data. They first identified the user's political leanings using a supervised classification model, and then looked at whether a user shared a political tweet or not, and whether those they retweeted had a reciprocal following (symmetric) or not (non-symmetric). Their measure of homophily was calculated as,

$$Homophily = \frac{outbound\ edges\ to\ users\ with\ the\ same\ political\ leanings}{total\ \ outbound\ edges} \quad (8)$$

Similarly, the study by Asher, Leston-Bandeira and Spaiser (2017) using UK Twitter data showed that more edges were seen within clusters of similar users than across clusters. Similarity in the study was based on attitudes towards an e-petition. Qualitatively, this suggested more people interacted within their own online communities than with other communities and therefore reinforcing their own political viewpoints (Asher, Leston-Bandeira and Spaiser, 2017). Analysing the meta-data of Twitter also found that a high percentage of tweets during a debate about Brexit were retweets rather than replies (Maynard *et al.*, 2017). Colleoni, Rozza and Arvidsson (2014) describe this as an 'echo chamber' environment and although is commonly seen in social media data, is a trait which garners conflicting support. One of the core questions of this research is: *What can we learn from participants' interactions during these digital discussions?* Therefore, homophily will be a good indicator of whether users in a network interact more with like-minded people.

Ideally one would like the participants of a discussion to come from many different backgrounds with different viewpoints in order to create a more balanced discussion. Getting people to interact with people they would not otherwise usually do, or listen to different points of view is a positive goal of parliamentary online engagement exercises (Mutz and Wojcieszak, 2009, p.49; Papacharissi, 2002, p.23; Carpini, Cook and Jacobs, 2004, p.335). This however, does not necessarily have to mean the opposite of homophily. Looking back to Figure 11, the *Tight Crowd* type shows a network with not only many edges within the clusters, but also across clusters suggesting the users are interacting with both like-minded users and one's from different communities. The *Community Cluster* type also shows this to a lesser extent.

Furthermore, at times homogeneity in groups can create an environment of misinformation or low quality information that does little to educate the members of wider political information (Conroy, Feezell and Guerrero, 2012). On the other hand, the evidence of homophily in political communication data is something some researchers feel is a positive trait of a network. The phenomenon of *Group Polarization* can be described as when people's views become more extreme when they are amongst others who also share their views (Sunstein, 2002). In other words, the existence of homophily creates more polarised peoples and this is amplified through the internet. He argues however, this phenomena on its own is not seen as a disadvantage, and only becomes a problem when the group's views become too extreme. Other researchers also argue that deliberative democracy works best when participants agree with each other. For example, Mutz and Wojcieszak (2009) found that discussion groups about civic, religious, ethnic, and political topics had the highest evidence of homophily compared to trivia or social topics for example. In this view, homophily is something to be valued in a participatory setting, and is encouraging to see when shown in a network graph.

Officials in the House of Commons are generally of the view that encouraging people from different backgrounds to participate is key to the effectiveness of their public engagement activities (Liaison Committee, 2019). They acknowledge that certain demographics including young people and those from low-income backgrounds are underrepresented compared to the usual older, more educated and well-off members of society (Liaison Committee, 2015; Walker, 2012).

3.5 Participant observation

As has been explained throughout this chapter, quantitative methods such as social network analysis and text mining make up the majority of the methodological framework of this thesis. However, thanks to the collaboration between myself and the House of Commons, participant observation has also been a valuable method of data collection and analysis.

Spradley (2016) notes that there are 5 criteria for successfully doing participant observation, "(1) simplicity, (2) accessibility, (3) unobtrusiveness, (4) permissibleness, and (5) frequently recurring activities." (p.52). During my time in this PhD I made regular visits to the House of Commons, specifically to work with the Digital Engagement Team. I began by simply observing them in their day-to-day work, attempting to understand how they approached digital engagement and what mechanisms they already had in place for the evaluation of the digital engagement activities. I spoke to each member of the team, as well as members of other teams such as Public Enquiries and Resources & Content Development teams who all shared the same office at the time. I wrote detailed notes on the informal conversations I had which helped me to reflect and understand the relationships between staff members and the over-arching structure of the House of Commons. This is explained in further detail in section 4.1. In this beginning stage I adhered to the first and second criteria of simplicity and accessibility (Spradley, 2016), and made an effort to be as unobtrusive as possible. I did not interfere with their daily activities more than asking them to talk me through how they would manage the digital discussions.

Having this insight into the team's responsibilities allowed me to clearly understand how engagement is currently evaluated, and the reasons for choosing certain methods over others. For example, they often measured the 'success' of an online engagement activity through the number of comments or views to a particular post because those were the key performance indicators (KPIs) the team was required to meet at the end of every month. Likewise, when a digital discussion had many comments, the team did not use any text mining approaches to analyse the text because they did not have those skills or access to software which could do it for them. As a result, they were required to manually ready through every comment (often reaching the thousands) and produce a summary of main themes and quotes to an MP within a short period of time. Therefore, when I came to assess how the evaluation of digital engagement activities could be better, I did so from an informed position having spent time among the people who would be using the solution I produced (in this case TheGist text analysis application). While my primary collaborating team was the Digital Engagement team, through frequent visits to Westminster and word-of-mouth, I also worked with the Web and Publication Unit and several select committees. Along with providing me access to additional data, this also gave me a further insight to how different teams and departments within the House of Commons operate, despite doing the same role of digital engagement.

There were also some disadvantages to having a collaborating partner for the PhD. The UK Parliament is naturally very risk averse and are very cautious with the type of data they collect and the methods of analysis they do as a result. For example, when analysing some of the data from digital discussions held on Facebook, the Digital Engagement team wanted to know the demographic information of their participants. As the data extracted from Facebook includes the names of users, I suggested analysing the sex distribution of the comments to see whether women expressed different sentiments or focussed on different sub-topics to men. This would be of interest to the team who wanted to understand who was participating in the discussions to ensure they were reaching as many different sections of society as possible. However, while the team was happy to use the comments from the participants, they did not want me to process any further information about the users, such as their names.

It is important to respect the concerns of the collaborating partner in any partnership and I have aimed to do so throughout this PhD. However, this thesis is also an independent piece of work which raises any failings or areas of improvement I observe. The research and analysis I complete remains thorough and unbiased.

Conclusion

This chapter has introduced and discussed different methods for analysing data from social media and participant observation, and how they relate to this research. Text mining is the primary method of data analysis and is used for processing and interpretation of comments left on Facebook, Twitter, and the Discourse platform. Different models for getting as much information from text as possible such as selecting the topics and sentiment from text are employed to help policy-makers and parliamentary staff to quickly assess the feelings and ideas of the public participating in their online engagement activities. In the same vein, social network analysis can help to uncover exactly who is participating in these activities and whether they are interacting with people outside of their social sphere or not. Using the Twitter dataset purchased by the House of Commons, the geo-location tag for some tweets may also allow me to understand where the participants are based, and perhaps inform the MPs of their constituents' issues. The choice of algorithm used in each method will largely depend on the input data and the final purposes of the project.

Many of the techniques explained in this chapter encompass text as a whole but using social media data is a specialised case as touched on regarding topic modelling and pre-processing in particular. Bringing the practice of parliamentary engagement to the digital arena allows for detailed insights into the feelings, behaviour, and patterns of the people participating with Parliament. In the following chapters of this thesis, examples of these methods are used to extract as much useful information and insights from text as possible. They will be used in the analysis of digital discussions on Facebook, Twitter, and Discourse platform experiments. The next chapter will introduce the various teams in the UK Parliament who work with the public and conduct digital engagement activities. I have worked with some of these teams by conducting participant observation as explained in this chapter.

Chapter 4 Evaluating engagement: the view from inside Parliament

This chapter focuses on understanding the data Parliament already possesses and analyses, and what it means for Parliament to be effective in their public engagement initiatives. In order to create a solid foundation to assess any initiatives, it is first important to understand exactly who within the organisation has remit over which areas of engagement, and their parameters of success. To date, a Parliament-wide organisational structure which identifies the teams dealing with (specifically online) engagement has not been developed. Parliaments are large and complex organisations and, as has been touched on in section 3.5 (and will be discussed in more detail in section 4.3) are not the most suitable institutions for the implementation of digital methods. However, as research shows, it is important to understand how the institution and its staff implements engagement initiative and what they expect from this; and, crucially, whether this differs from service to service. As such, there exists an environment in which several teams may have a similar goal, albeit not working together due to a lack of communication or awareness. There can be many reasons for this including different priorities for different teams, as well as different levels of expertise and influence within the institution. This diversity of goals and understanding of public engagement clearly has an impact on the way in which they are evaluated. It can lead, for instance to differences in the understanding the interpretation of what 'effectiveness' means for each team within Parliament, and for each type of activity. This chapter introduces an organigram of the teams dealing with digital public engagement (as characterised by the Public Engagement Spectrum outlined in Figure 4) with the aim of identifying the organisational silos that exist within the institution today, to help me better understand how tools to evaluate public engagement can be developed.

The structure of this chapter will be as follows: firstly I introduce an organigram of the teams that deal with public engagement in parliament. I then take a closer look at how these teams implement and evaluate public engagement to help me understand their interpretation of an effective engagement activity. This will in turn enable me to develop criteria to assess these public engagement activities, before exploring the challenges associated with introducing new technology to Parliament.

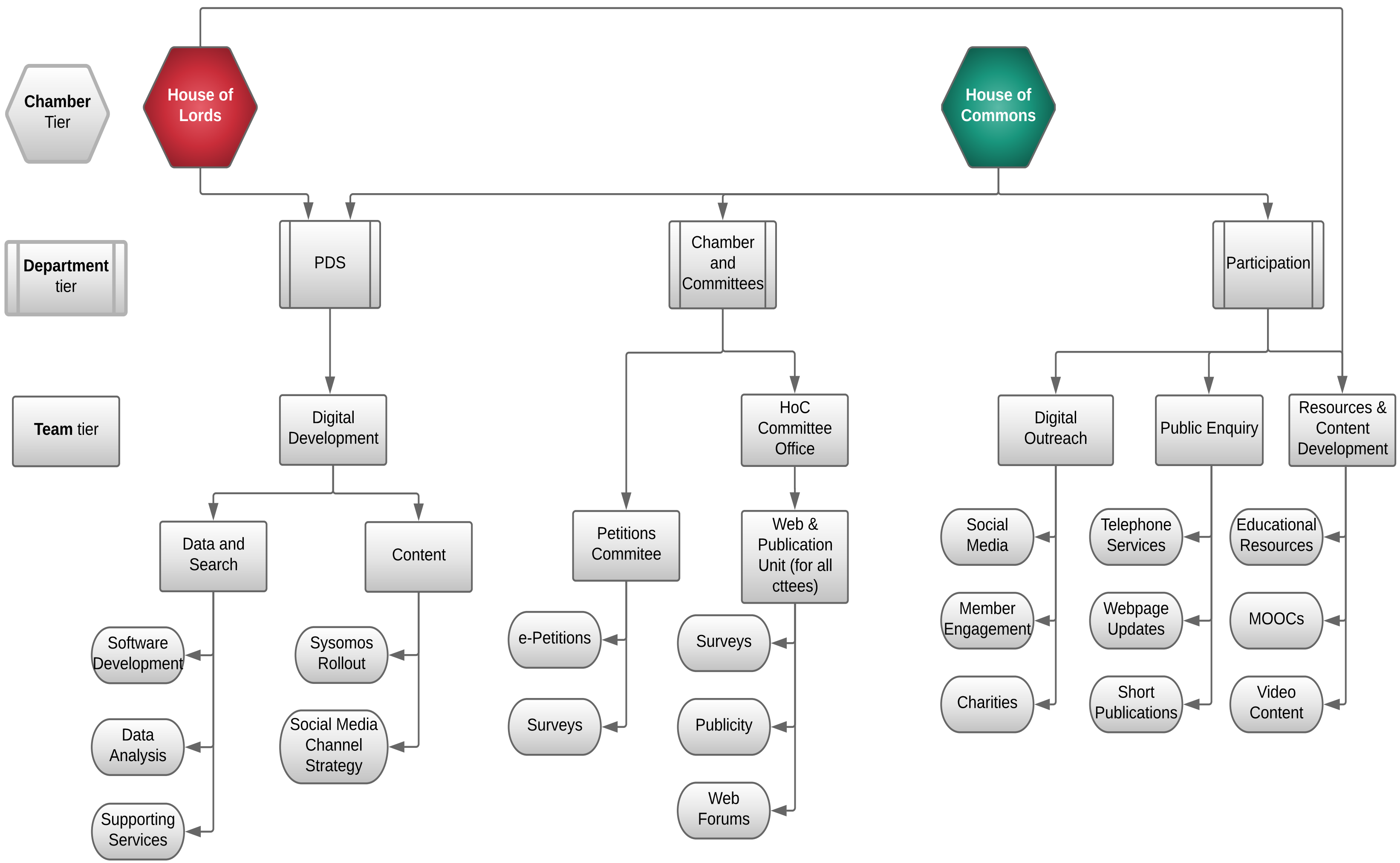
4.1 Organisational structure of digital engagement in the UK Parliament and their engagement activities

This section introduces the current organisational structure of the teams responsible for online engagement within the UK Parliament. It explains their main roles relating to digital engagement and how they correspond to the dimensions of engagement explained in Figure 4. This is an important first step in understanding the workings of engagement in the UK Parliament for several reasons. Firstly, 0 demonstrated how public engagement can cross different dimensions ranging from those activities which seek to inform or educate the public to those which seek to obtain views and enter into a discussion with the public (Arnstein, 1969; Lenihan, 2008). Depending on the area of Parliament a team is based in, they can be focussing on one or more of these dimensions at a time within a single activity; different teams may also be focusing on very different types of purposes and aims expected from their engagement activities. Therefore, mapping their organisational structure helps to understand which types of engagement a team is most likely to undertake and what objectives they are likely to favour. Secondly, through understanding the type of engagement, I can also begin to explore the measures currently put in place by different teams to assess the effectiveness of their activities (Coleman and Firmstone, 2014). By categorising teams into departmental areas, I can develop a better understanding of how the different teams are evaluating their own engagement activities, and what different teams most value. I can also see if these measures of effectiveness are determined by the type of engagement activity or moreso by the similarity of activities done by other teams in the same area.

There are two chambers in the UK Parliament, the democratically elected House of Commons and the appointed House of Lords. The principal teams responsible for online engagement are unicameral and based primarily in the House of Commons with the exception of the Parliamentary Digital Service PDS which is bicameral[10]. There are over 400[11] employees in PDS which is split into several sectors and sub teams. The Digital Development sector includes 5 different sub-teams, two of which are Data and Search and Content (Parliamentary Digital Service, 2017). These two teams handle the development of the new website, creating new content for the current website, and introducing new analytics tools to name a few. The content team also works closely with several other teams in Parliament when required.

[10] This was true as of my participant observation during 2017/2018 but may have since changed.
[11] As of 2018, Personal communication, Westminster

Figure 13: Organigram of teams dealing with digital engagement in the UK Parliament



Within the unicameral departments, Chamber and Committees and Participation are the two most closely linked to digital engagement. The Chamber and Committees department includes (amongst other) the Web and Publication Unit (WPU) which sits under the House of Commons Committee Office team. The WPU is responsible for the digital public engagement of all of the select committees in the House of Commons, and organises any campaigns, surveys, or any other publicity a committee would like to promote for one of their inquiries (which consist the main bulk of committees' work). An exception is the Petitions Committee who have dedicated Public Engagement Specialists employed on their team to conduct their public engagement activities. These roles unique to this committee provide it with more expertise in terms of the understanding of engagement, meaning they rely less on the WPU than other committees may do.

The Participation department consists of 3 teams dealing directly with online engagement; Digital Outreach, Public Inquiries, and Resources & Content Development. The Digital Outreach team (now renamed to Digital Engagement) conducts the Digital Debates on Facebook, runs the @HouseofCommons Twitter account, and uses other established online communities such as Money Saving Expert and The Student Room to gather opinions from the public. They liaise with different specialised charities around the country, so they have a ready-made community to turn to when an MP chooses to hold an online debate. They also work closely with the Petitions Committee, to have a discussion on the House of Commons Facebook page to gather the public's views on specific petitions. The Public Enquiries team is the first port-of-call for enquiries from the public and the first engagement opportunity many will have with the institution besides the website. They receive around 15,000 enquiries a year with 2/3rds of these being over the phone[12]. The team also maintain some of the parliament.uk webpages and produce short public publications for the House of Commons. The Resources & Content Development team is bicameral and are responsible for creating online and offline content in the form of educational publications, video content, podcasts, and MOOCs.

The spectrum of engagement introduced in section 2.2 separates engagement into activities that require the input of the public and those that do not. Focussing first on the latter, this relates to those activities which aim to inform the public about Parliament by educating, disseminating information, or through outreach and openness events. The former, being activities which require input from the public are characterised by consultation and discussion-type activities. However, this spectrum only identifies the activities themselves without making any link to the actual people behind the organisation of the activities. Without an understanding of the motivations of the teams running the engagement sessions it will be difficult to understand how they can evaluate their digital engagement. The teams identified by the organigram in Figure 13 can be categorised by the dimension of engagement their activities fall under as described in Table 2. This reveals that while having distinct responsibilities, each team can occupy multiple areas of the spectrum and, in doing so, find similarities in practice with other teams across parliament.

Teams with duties lying in the Information section are Digital Outreach and Public Enquiries. Both of these sit under the umbrella of Participation and use Twitter and the parliament.uk website to disseminate information to the public. Public Enquiries also use the phone lines and email to answer questions or point people in the right direction. The Resources and Content Development team and the Content team both sit within the Education dimension. As these are the only two bicameral teams in the organigram, it suggests this dimension covers the whole of the Parliament more than any other dimension. These teams use online resources for teachers and lecturers, as well as online courses which have proved extremely popular with the public, especially when

[12] Personal communication, Westminster, 28 March 2018

disseminated through Facebook. Data which I have analysed from Facebook shows the 'Introduction to Parliament' MOOC was one of the most popular posts posted by the House of Commons account between April 2017 and May 2018. Once again, in these cases the source of information lies with Parliament and the direction is one-way.

The other branch of the spectrum is concerned with activities that have the public as the source of information rather than the Parliament, and in this way aim to encourage a two-way discussion. The Petitions Committee, WPU, and Digital Outreach teams all sit under this branch. Each of these reaches out to the public to gather their views on a certain matter and relay that information back to MPs and officials. In this organigram, the e-Petitions are a special case due to the actor initiating the activity. Although the information comes from the public, most consultation activities above are initiated by Parliament; if a Select Committee needs information for an inquiry, they will themselves create and share a survey asking the public a set of questions; if the Digital Outreach team have a topic suitable for Facebook they will create a digital discussion card again asking the public's views on that topic. However, the e-petitions are created and owned by the public, and simply hosted on the petitions website. The petition creator is responsible for sharing and making others aware of the petition and bringing a particular issue to Parliament's attention. Only once enough members of the public have added their support through signatures does Parliament take action. Lastly, a similar list of teams lies within the Participation category which involves information flowing from both actors (the public and the institution). In these cases, the institution is an active participant in the discussion, not simply asking for or giving information.

Table 2: Teams categorised by dimension of engagement

Dimension / Category	Inform the Public		Encourage Participation	
	Information	**Education**	**Consultation**	**Discussion**
Team	• Digital Outreach • Public Enquiries	• Content • Resources and Content Development	• Petitions Committee • WPU • Digital Outreach	• WPU • Digital Outreach
Channel	• Twitter • Website • Facebook	• MOOC • Resources • Video	• Surveys • Forums • Inquiries • E-Petitions	• External online communities • Facebook
Possible effectiveness measures	• Wide geographical reach • Communication across network communities • High follower numbers	• Video views • High course completions	• Positive or negative sentiments • Signature count • High comment length • Diverse participant characteristics • Range of topics	

4.2 Understanding effectiveness from the perspective of different organisational units

Understanding what it means to be effective requires the incorporation of many different factors. In the case of the UK Parliament, just as public engagement does not fall into solely one dimension (Walker, 2012; Rowe and Frewer, 2005), neither does the interpretation of effectiveness. This chapter argues that in order to understand and measure how effective a public engagement activity has been, it is not only necessary to understand which dimension of engagement the activity falls under, but also which department of Parliament is leading the activity. Therefore, in order to answer the question of effectiveness, a clear understanding of the underlying processes and the steps taken leading up to the activity is vital.

Earlier, Table 2 showed that one team is not restricted to one area of engagement, providing also a 'possible effectiveness measures' row by which the dimensions of engagement can be categorised. Focussing on the activities within the *Inform the public* branch, the main aim of these services is to reach as many people as possible. With this in mind, an effective information dissemination activity through the website could be characterised as the number of page visits within a certain time period, the number of views of a debate on the parliament.tv website, or the reach and follower numbers for a Twitter account. These measures all contribute to an understanding of how receptive the public is to an activity. For example, the Introduction to Parliament online learning course run by the Resources & Content Development team recorded 10,000 users during its 3-week run in 2018[13]. This was seen as a successful engagement activity and the high number of users was used to justify the renewal of the course platform license for another 12 months.

These are all very quantitative measuring frequencies and counts but are sufficient in providing enough insight into how many people have been reached as this is the primary aim of these activities. On the other hand, for activities in the consultation and discussion dimensions of the public engagement spectrum, general summary analysis of follower counts and page views may not be sufficient to understand the specific feelings and opinions of the public, and therefore more detailed measures are needed. For these activities, measures of effectiveness include those counts, along with measures which are more in-depth and analyse the characteristics of those who have participated, how substantial their submissions are (comment length) and how they felt about the topic. An example of this lies with Twitter. Summary statistics such as number of likes and retweets are useful for knowing the reach of a particular post, however the quality as well of quantity of posts are even more important. A discussion with few comments may be seen as rather unsuccessful, as there would not be many insights and opinions that the MP could use to inform their speech. Reaching a large number of people is a factor into why an MP would choose to have a Facebook debate in the first place, so getting many comments suggests a large portion of people are interested and have had an opportunity for their opinions to be heard. Likewise, a discussion with thousands of argumentative comments would be equally as unhelpful for an MP to extract meaningful stories and anecdotes without some form of automatic analysis. On the other hand, in their study of Facebook and Twitter use of Spanish local governments Haro-de-Rosario, Sáez-Martín and del Carmen Caba-Pérez (2018) found that comments expressing a more negative sentiment led to more enagement overall because dissatisfied users were more likely to to complain and in turn, engage. Therefore, quality as well as quantity are important factors when understanding the effectiveness of an engagement activity.

Although two different teams may operate within the same dimension of engagement, they can still have very different processes due to their position in Parliament and their remit. Due to their differing priorities, the select committees and the Digital Engagement team have different approaches when conducting engagement sessions. Digital Engagement sessions are more heavily influenced by the business of Parliament, and whichever issues MPs are debating at the time. This will often be through the digital debates mentioned in the previous section which focus around an issue close to the MP's constituency, a topic an MP is working on, or perhaps a popular e-petition.

[13] Personal communication, Westminster, 28th March 2018

These are often short consultations on Facebook or through surveys lasting around one week and occasionally the MP will participate. The Digital Engagement team also uses established online communities (The Student Room, 2020). In the case of the Mumsnet forum, an example of one of these established online communities, the Digital Engagement team regularly conduct consultations by asking members of the forum to give their views on specific topics being debated by MPs. At the end of 2019, the Digital Engagement team produced a summary of all the input the Mumsnet members had provided throughout that year and exactly how their comments had been valued and taken into account (House of Commons, 2019b). This was shared with the Mumsnet community to thank them for participating, and closed the feedback loop between the public and the institution which is a crucial aspect of many public engagement activities which often gets overlooked (Smith, 2009; Coleman and Gotze, 2001). On the other hand, select committees have a range of different points on which they can consult the public, and these can fall into different dimensions of engagement as explored in section 2.6. This can be informing the public about the launch of a new inquiry, distributing a survey, or holding an online consultation session encouraging the public to participate and give their views. In this way, select committees have a wider scope and opportunity for different types of engagement, that are not as closely dictated by parliamentary business as the Digital Outreach team.

The progression of engagement is another form of effectiveness to be measured. For example, following a parliamentary account on social media or accessing the parliament.uk website could be a starting point for a member of the public interacting with the institution. Through this they may see details of an e-petition they would like to sign, then contribute to an inquiry as a result of the e-petition, then go to an evidence session at Westminster or watch it online. Each of these activities would be handled by different teams within Parliament, who would each record the interaction as a separate event and have their own success measures. For example, “Joe” could see an interesting petition shared on the UK House of Commons Twitter page. He signs the petition and the Petitions Committee would record his signature on their database. Another committee may then hold an inquiry asking for written evidence which “Joe” could submit to them. “Joe” may then want to watch the inquiry’s evidence session online which would be recorded as views on the parliament.tv website. He may then wish to visit Parliament to see the chamber for himself which would be handled by the visitor’s centre. Each of these teams working independently of each other will record him as a separate participant to their activities without knowing from where or how he came to know about it. While this contributes to the key performance indicators of a team’s quarterly or annual targets, an understanding of how the action by a member of public is part of a larger process of building a relationship with the institution is lost.

In the case illustrated above, rather than measure the effectiveness of these activities individually, there could be a cumulative measurement of the end result. This would allow teams to understand how the success of their activities is impacted by other teams, and track what encourages the public to repeatedly engage with Parliament. This progression of engagement is something the WPU team is particularly interested in[14], as this would allow them to seek new communities of people who may provide valuable evidence to a future select committee inquiry. Recognising this user journey is not guaranteed as it can be difficult to track direct relation of one activity to another, and correlation of two activities does not always mean one caused the other. For example, creating a new online course and receiving more comments on a Facebook post; the people who completed the new online course may be more knowledgeable on how they can interact

14 Personal communication, Westminster, 3rd May 2018

with Parliament and then choose to respond to a Digital Debate on Facebook, but it is difficult to prove this link. Nevertheless, rather than being restricted to only one dimension of engagement, citizens can float from one branch of the spectrum to another going from information to consultation back to openness.

One example of this is the case of the Digital Outreach team who work closely with the Petitions Committee, as the majority of the traffic to the e-petitions website comes from Facebook. Many of the digital discussion cards held on the UK House of Commons Facebook page are e-petitions which have reached the 100,000 signature threshold and are waiting for a debate. One digital discussion led to a participant being interviewed by the Digital Engagement team over the phone and invited into Westminster to view the debate[15]. This shows how a participant went from viewing an e-petition to participating in a digital discussion to physically visiting Westminster to watch a debate in person. Their engagement with the UK Parliament covered many different spectrums including consultation, education, and openness.

This inter-connection between different dimensions is one reason why the two branches of the spectrum are at equal levels in the conceptualisation, rather than as a ladder signifying sequential steps or implying that one area of engagement is more important than another as in (Arnstein, 1969) . The two sides work hand-in-hand and often a particular activity may overlap several dimensions. Just as the same team can occupy different positions on the engagement spectrum, so too can the public. There is also evidence to suggest that the ways the teams work in practice does mirror this conceptualisation of the engagement spectrum. At the beginning of this collaboration in 2017, the Digital Outreach team's work with social media was informally separate, depending on the type of content they were sharing. Daily business of the House of Commons and questions asked in the chamber were posted to Twitter. These types of posts are intended to inform the public of daily news rather than elicit particular responses. Therefore, an analysis of how many people were reached, as well as likes and retweets are suitable metrics on which to measure the effectiveness of a post. Conversely, discussion cards posted to Facebook directly asking the public's views on a particular topic raised by an MP were designed specifically to stimulate a response from the public. However, by 2019, this separation of duties based on engagement within the team was made formal when the team was split. Those updating Twitter were relocated to the Communications department while those working on Facebook remained under the umbrella of Participations.

Deciding which of quantity (counts) vs quality (in-depth) prevails in terms of effectiveness is also dependent on the team leading the activity. A situation where quality could prevail lies with committee inquiries. For the WPU, a small set of people who have a true experience of the inquiry topic and who have been affected by it in their day-to-day lives would provide a much better set of inquiry evidence than a large group of people who do not have first-hand experience or are not directly affected by the issue. Conversely e-petitions provide an example of quantity being very important, where to be considered for a debate, it must reach at least 100,000 signatures. Facebook Digital Debates held by the Digital Outreach team also value the number of comments posted as a key performance indicator to measure the success of the activity. For this reason, understanding both the team initiating the activity and the dimension of engagement it falls under is crucial to its accurate evaluation and assessment.

With all this in mind, I must remember one of the motivations for holding a discussion online in the first place. Gathering opinions from a wide range of people from different backgrounds and different parts of the country is not easy to do in person. Therefore, using digital

[15] Personal communication, Westminster

engagement by definition implies receiving a large volume of data and is one of the obstacles for teams conducting large-scale online engagement activities.

4.3 Introducing new tech to Parliament

There are many considerations when working with the UK Parliament, especially regarding new technologies. The UK Parliament has had a contentious relationship with embracing technologies in its past with radio and television broadcasting being introduced only in 1978 and 1989 respectively (Norton, 2005). Nowadays, they are more open to using tech which can improve or facilitate their day-to-day tasks. With this in mind, there still exists a difficulty in terms of the risk-averse nature of Parliament, especially when dealing with new technology. Data privacy, security issues and logistical difficulties all need to be considered. This is particularly important for an institution which values its reputation as highly as the UK Parliament. They are always aware that they cannot be seen to be using tools which could introduce biases of any kind into their work. They are notoriously risk averse because unlike political parties and MPs, Parliament must present an unbiased view and remain neutral, while also representing the beliefs of all MPs and parties who have been elected. Therefore, having one single image to portray themselves can be difficult (Kelso, 2007) and as Leston-Bandeira (2012, p. 422) notes parliaments can be very vulnerable to negative effects on their image "which has consequences into the development of public engagement activities" . While these consequences do indeed affect the engagement activities, they also affect the methodological and technical aspects of engagement, namely what types of analysis are completed and what kinds of tools are used to achieve this.

For example, following the news of Cambridge Analytica's involvement in the 2016 US presidential election and the 2016 Brexit vote (Persily, 2017), conducting analysis on Parliament's behalf using Facebook data became more difficult. It was vital to emphasize the differences in Cambridge Analytica's work and my own to parliamentary officials involved in the collaboration, to make clear only information which is visible in the public sphere is being analysed in accordance with GDPR. Fortunately, any analysis of data was authorised under the 'legitimate interest' of GDPR guidelines[16] and obtained a full ethical review clearance from the University of Leeds[17]. Had this ethical approval not been obtained and the senior managers of the participation department not been reassured of its validity, they may well have restricted my access to the House of Commons Facebook account or changed the nature of the project entirely.

There also exist difficulties in how to implement new tools into the UK Parliament and work alongside the procurement process. During 2018, there was a lengthy and costly process of implementing a new social media listening tool, Sysomos, across several large teams within the House of Commons and House of Lords[18]. This was intended to improve understanding of the many social media accounts held by different teams and to filter large volumes of tweets for better understanding of trending topics. It could also be used to schedule social media posts and descriptive analysis of account performance. However, the tool ended up not being used by the intended teams for the right purpose.

There were numerous reasons for this, but one major flaw was its lack of specificity to each team and the large number of features which led to confusion and eventual frustration with the

[16] Personal communication, Westminster, 1st May 2018
[17] Reference AREA 17-157
[18] Personal communication, Westminster, 29 November 2018

tool. Unfortunately, the staff members who already had limited time resources could not justify taking further time to use a tool which was not guaranteed to reduce their workload significantly. Although this tool was not adopted widely in Parliament, important lessons were learned on the value of having a tool created specifically for the teams using it. Due to this, I develop a text analysis tool, TheGist, which does precisely what is needed by the House of Commons. Further details on this tool are explored in Chapter 7.

Despite TheGist being created solely for the UK Parliament and their involvement in its development from the beginning, there have been some difficulties with using the tool. TheGist was created using the R Shiny platform which allows building of web applications using R (RStudio Inc., 2013). This allows the application to be fully interactive and includes an option for deployment using the Shinyapps.io cloud service. At its most basic, the application can be accessed by anyone who has the URL through a web browser, however this is not ideal due to the potentially sensitive nature of the data uploaded to the application. There are several options for deployment including ones which limit the number of active hours an application can run for, and ones which include authentication of the application to make it password-protected (Shinyapps.io, 2020). A password-protected application ensures the team can use TheGist while still adhering to data privacy policies within Parliament, however this option comes at an additional cost of USD $99 per month. This extra cost was not budgeted for by the team so the use of the application is delayed while arrangements are made to pay for this.

However, introducing new technology to Parliament is not always a difficult task. In 2015 e-Petitions were introduced following a report by the Digital Democracy Commission (Commission, 2015) to provide the public with another way to interact with the Parliament. However it is important to note that the concept of e-petitions had previously been raised by several influential reports over the years including Modernisation Committee (2004) and (Reform Committee, 2009). This involved creating a new Select Committee for Petitions and a new website run by the Government Digital Service to manage the incoming petitions (Petitions Committee, 2015). This was well adopted by the Parliament (Petitions Committee, 2016a) and continues to be a valuable and approachable channel for the public to raise important issues or initiate a dialogue with the institution.

During May and June 2019, I conducted demonstration tests with inquiries from three select committees trialling a new online discussion platform called Discourse (Discourse.org, 2019). While there were some restrictions about which digital platform I could use due to location of servers and cost, once Discourse was chosen, there was much enthusiasm from the committee staff involved. During the demonstration tests, they allowed me to have full control over the platform set-up and management which was done remotely. I gave regular updates to the committees throughout the duration of the online discussions and provided interactive summary reports on the comments posted by participants for each of the inquiry topics. The use of this new platform in Parliament was cited positively in several of the committee reports (Environment Food and Rural Affairs Committee, 2019; Transport Committee, 2019), as well as an evaluation of the effectiveness of the select committee system by the Liaison Committee in the same year (Liaison Committee, 2019). In the case of the Environmental Audit Committee, a new question was asked to the Parliamentary Under Secretary of State for Rural Affairs and Biosecurity as a direct result of the responses collected during the demonstration test with the new online platform. There is clearly plenty of appetite to try new things and parliamentary staff are happy to do so.

These three committees showed that they are happy to embrace new technologies in their work, and other committees also showed an interest in using Discourse in the same way. Despite the enthusiasm from parliamentary staff, the demonstration tests would have been much harder to implement had I not first discussed them with the then-Clerk of Committees who instructed the committees to work with us. So, what usually gets in the way of technology being introduced and embraced in Parliament is the structure and the internal processes of the institution. This then perpetuates the risk averse nature and the slow decision-making. Conversely, when attempting to recreate the experiments in the Digital Engagement team, there was more hesitance in using the tool due to data privacy and storage concerns. The House of Commons has a strict data protection policy which includes restrictions over exactly what data can be stored from the public and how it can be used (House of Commons, 2019a). Where the select committee teams had little problem with this, the Digital Engagement team were explicit in how this should be handled before proceeding with the tool. So it is clear that Parliament has some difficulties and hesitancies when introducing new technology, however as Kelso (2007, p. 372) remarks "to acknowledge change only when it happens in a `revolutionary' way is to fundamentally miss the evolutionary character of Parliament, and to miss the incremental accumulation of important changes over time". While introducing a new platform and using it for an inquiry evidence session may not appear as a substantial feat, it all contributes to a better understanding of what kinds of tools can be used to facilitate engagement and its evaluation and encourage Parliament to be more open to their uses.

However, the most vital part of introducing new technology is having the processes to manage it and ensuring its use has been carefully thought out. Sysomos was not specific enough to each teams' purposes which led to it not being used in the intended way. With TheGist application, the budgeting for using a web-based application securely was not accounted for leading to delays in its deployment. And with Discourse, different areas of Parliament put varying importance on the data privacy restrictions due to the influence of senior parliamentarians leading to only one department (select committees) benefitting from the platforms' use.

Conclusion

This chapter has explored how the different service teams in the UK Parliament manage their digital engagement activities; what they are responsible for and, correspondingly, how they evaluate their activity. I explored what it means to have an 'effective' digital engagement activity for different teams within the UK Parliament, that one team's activities can fall into a range of different dimensions of engagement as categorised by the spectrum in Chapter 2 (Figure 4), and as such requires multiple methods of assessment. Dividing teams based on the type of engagement they undertake is already in practice for example with the Digital Engagement team's separation of tasks based on information dissemination and consultation activities. This is an important aspect of the UK Parliament that must be clarified before any suggestions on evaluation techniques can be made. I must understand how and why certain engagement activities are conducted the way they are and how the purpose and aims of different teams influences how they will ultimately evaluate an engagement activity's success or failure. While facilitating the evaluation of digital engagement activities remains a core focus of this thesis, it is also important to recognise the internal processes and risk-averse nature of Parliament which can restrict the introduction of new technology. Evaluation of large-scale digital engagement activities requires the use of novel platforms and computational methods which Parliament has acknowledged but still seems reluctant to embrace (for example their use of the Sysomos platform).

With a better understanding of the types of engagement being done in Parliament, the next chapter begins to explore some of the activities controlled primarily by the Digital Engagement team on Facebook and Twitter. These activities lie within the information dimension of the engagement spectrum and as Table 2 suggests, can be measured through reach of different social media accounts and social network analysis.

Chapter 5 Assessing Parliament's information outreach

The previous chapters have established the different dimensions of public engagement (0) who is leading the digital engagement activities in Parliament and what they deem to be an effective activity (Chapter 4). In this chapter, I begin to put the evaluation metrics of online public engagement activities I explored in Chapter 3 by examining some real-world data from parliamentary digital engagement activities on Twitter and Facebook. These activities aim to inform the public about events and daily business of Parliament

As described in section 3.1 the Twitter data was purchased through a third-party company hired by the House of Commons on 28th March 2018. The Facebook network data was collected through the Facebook Netvizz application in December 2015, and the Facebook post data through API web scraping in May 2018. Since 2018 there have been developments, however the analysis in this chapter is accurate to this date. These are assessed through the examples of evaluation methods of activities which aim to inform the public (introduced in Table 2) such as evaluation of geographic reach, numbers of followers and communication across social networks.

Sections 5.1 and 5.2 focus on the UK Parliament's accounts on Twitter which encompass information dissemination dimension of engagement and are measured through more quantitative methods using Twitter, while section 5.3 focuses on analysis of Facebook posts which are made by the UK House of Commons account run by the Digital Engagement team. Unlike what will be explored in future chapters, these Facebook posts are not made by citizens but rather by the UK Parliament staff, and as such are still categorised as information dissemination.

5.1 Reach of Parliament

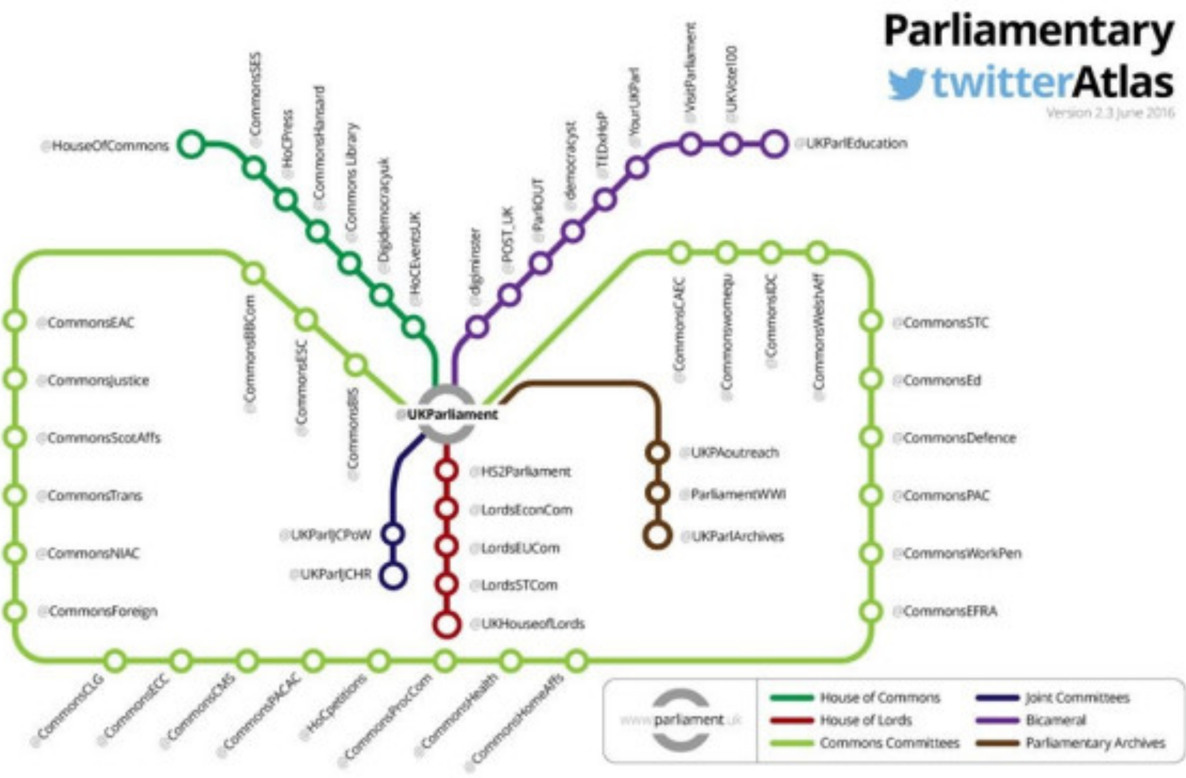
Since its creation in 2006[19] , Twitter has been used by many public institutions to provide information to the public (Inter-Parliamentary Union, 2016). Its unique 280[20] character limit on posts made it an attractive channel for the public to use, but also forced Parliament to take a different approach to explaining what they do. Where long, jargon-heavy publications were commonplace among MPs and staff, Twitter encouraged staff to think about conveying their work in a simpler way that was more concise and accessible to the general public. This new way of communicating also helped the Parliament become more open as it requires them to change how they interact with the public and creates new avenues for engagement. The use of Twitter has been primarily by individual political figures with a specific agenda (for example to communicate with the public and improve their popularity with the end goal of getting votes for an election (Cozma, 2013),) or political parties with a specific agenda they want to publicise. While this is also the case with British politicians on Twitter, the UK Parliament and select committees have made use of social media to put forward a non-partisan and impartial view of the issues facing the public. As previously raised in section 3.5, the UK Parliament has a specific role in representing all views of MPs and parties who hold a seat. As such, the content and purpose of their social media presence is to inform the public of their work rather to persuade.

The UK Parliament began to use Twitter in 2009, starting with only the *@UKParliament* account managed by the first Digital Engagement Manager. Ten years later, this has grown to over 50 different accounts, each proving information on specific areas of Parliament from select

[19] https://twitter.com/jack/status/20
[20] Character limit was 140 at the time of data collection and analysis

committees to education centres. As public engagement became a core task of select committees in 2012 (Committee, 2015), following the success of the *@UKParliament* account, they started creating accounts to promote their work in scrutinising the government. Figure 14 provides an overview of the Twitter accounts owned and run by various teams in the UK Parliament. The data analysed in this section includes accounts on the light green, dark green, red, and purple lines representing the House of Commons, House of Lords, Bicameral, and Committees.

Figure 14: Parliamentary Twitter Atlas (Anikka Weerasinghe and Ramshaw, 2018)



As I show in Table 3, data correct to March 2018 shows Parliament's Twitter accounts have a reach of almost 2 million users which is only 3% of the total UK Population. However, this accounts for around 15% of the 13.7 million Twitter users in the UK in 2018 (Statista, 2019b) or, 1 in every 7 UK Twitter users follows a non-partisan parliamentary account. This also aligns with the Hansard Society Audit of Political Engagement, which found 12% of their respondents followed a politician or political party on social media (Hansard Society, 2018). This information does not include the non-partisan accounts of the UK Parliament but does provide some insight into the following habits and representativeness of Twitter. This is still a small overall percentage, however research from the Pew Research Centre found that even a small sample of Twitter users can have some common characteristics to the wider population (Pew Research Centre, 2019). For example, they found that American Twitter users differ from the population on demographics as they are generally younger and more democrat than the broader population. However, in some other subjects their views align well with most Americans.

Our Twitter dataset contains 48 of Parliament's Twitter accounts; 2 covering each of the chambers - Commons and Lords; 6 covering the wider parliament; 33 covering different House of

Commons select committees; 5 covering different House of Lords select committees; and 2 covering joint select committees. These are tabulated below (see Table 3).

Table 3: Aggregated overview of Parliament's Twitter accounts (March 2018)

Area of Parliament	**Number of Accounts**	**Number of Followers**	**% of Followers**
HoC Select Committees	33	239,829	13%
UK Parliament-wide	6	1,098,238	59%
HoL Select Committees	5	73,783	4%
Joint Select Committees	2	6,790	>1%
HoC and HoL Accounts	2	452,896	24%
Total	**48**	**1,871,238**	**100%**

As can be seen from Table 3, cumulatively Parliament's Twitter accounts have a reach of almost 2 million Twitter followers and the majority (59%) of these represent those following the whole UK Parliament – specifically the *@UKParliament* account. Conversely, the accounts representing the joint select committees of the *@humanrightsctte* and *@jointctteNSS* have the smallest reach with less than 1% of the total share of followers. However, it is important to note that there are very few joint committees and they are very specialised, so are unlikely to have as much reach as other accounts.

Regarding representation of parliamentary departments, the select committees in the House of Commons have the largest share of the overall parliamentary Twitter accounts coverage. This is to be expected as they are the core of parliamentary activity and one of their main aims is to involve the public in their scrutiny of government (Liaison Committee, 2015). Therefore, creating an online presence through which they can interact with the public and provide information about their latest inquiries and reports is an important part of their work.

However, although select committees have the most individual accounts, their total following is smaller than the *@HouseofCommons* and *@UKHouseofLords* accounts, and dwarfed by the *@UKParliament* account. The reasons for this can in part be explained by the length of time each account has been active. The *@UKParliament* was created in May 2007, just under a year after the social networking site was launched[21]. Whereas, the Petitions Committee account for example was only created in July 2015 when the committee was itself created (Petitions Committee, 2015). Many other select committee accounts were created post 2012 – 5 years after the UK Parliament account, therefore it is difficult to make a direct comparison of influence based on follower numbers alone. However, the discrepancy in follower numbers can also be related to the subject matter of the accounts. The *@UKParliament* account claims to be "Keeping an eye on government, debating laws, approving taxes." which covers a large range of topics. The account also frequently retweets posts from other parliamentary accounts so its followers receive an all-round coverage of bicameral parliamentary news without necessarily having to follow each account separately. Of all of the Twitter accounts analysed, this is the most general and broadest account which would most likely attract both those with little knowledge of the institution and those who have a decent level of knowledge and would just like to keep updated. On the other hand, the select committees have a much more specific remit on which they tweet about, and as a result attract a smaller but arguably more involved set of followers.

[21] https://twitter.com/jack/status/20

This distinction between areas of Parliament also leads me to revisit the discussion of quantity vs. quality and to an understanding of quality in terms of Twitter followers. The quantity of followers for the *@UKParliament* account is very high, however the quality (in terms of them being legitimate people or bots) of those followers may not be high. When conducting inquiries, select committees aim to have a subset of the population with a specific interest, expertise, or lived-knowledge of a certain topic[22]. As some of these inquiries may be very specific, at times only a relatively small number of people will be suitable to contact. Therefore, having a small number of people with specific expertise is more valuable than a large number of people with little to no knowledge of the topic. For example, the Northern Ireland Affairs committee launched an inquiry into the impact of Brexit on the fishing industry in Northern Ireland. The select committee engagement team travelled to a fishing port in Northern Ireland to interview a few of the local fisherman as they knew they had a lived experience of the inquiry issue (Northern Ireland Affairs Committee, 2018, p. 10). This also relates to the different interpretations of an 'effective' public engagement activity – often depending on the initiator of the activity as discussed in section 4.2.

Another factor to be aware of is that a large number of followers does not mean they are all legitimate. Figure 15 ranks the accounts in decreasing order of average reputation score of their followers. This reputation score is calculated as the proportion of number of followers to number of those following. Therefore, Twitter users who have many Twitter followers but few Twitter friends, i.e. Twitter users that they themselves follow, would have a higher reputation. The equation is listed below:

$$Reputation\ Score\ = \frac{number\ of\ followers}{number\ of\ followers + number\ of\ friends}$$

Those users with a score closer to 0 are often interpreted as bots, as this shows they follow a large number of other users but have very few followers themselves (Chu *et al.*, 2012; Gurajala *et al.*, 2016). They find that bots have a reputation score of between 0 and 0.3 while most human accounts have a reputation score of over 0.5 (Chu *et al.*, 2012). Similar research into the reputation or influence of twitter user accounts define several measures of influence including in-degree influence which is described as where "the number of followers of a user directly indicates the size of the audience for that user" (Cha *et al.*, 2010). Research finds that users with large in-degrees signifying that they have many followers are most likely public figures or media sources. These users are contrasting to bots and would have reputation scores close to 1. On the other hand, some researchers suggest the reputation or influence of an account should be domain specific in order to "determine who is more consequential in this particular context" (González-Bailón, Borge-Holthoefer and Moreno, 2013). They compare a ratio of the number of messages received and sent by an account to the ratio of users an account is following and is friends with. In this way, they identify four groups of users; *influentials, hidden influentials, broadcasters*, and *common users*. Those users identified as *influentials* are mentioned most often and can be likened to users with a high reputation score. The common users have a higher ratio of following to followers and send more messages. However, this work is focussed on identifying user groups of protest diffusion and how they impacted the growth of the protest, and as such has a very specific domain rather than the varied subjects of the parliamentary Twitter accounts.

Most accounts have reputation scores in the range of 0.28 – 0.33 with an average of 0.30. As the *@UKParliament* account has many more followers than any of the other accounts, it is

[22] Personal communication, Westminster, 28 March 2018

expected that their followers have a relatively low average reputation score of 0.18. However, this is not the lowest score as the *@LordsSTCom* (HoL Science and Technology Committee) account followers have an even lower average score of 0.14. In fact, of the six accounts linked with the House of Lords, 5 of them fall in the 10 accounts with the lowest follower reputation scores. Only the *@LordsAICom* (HoL Artificial Intelligence Committee) is outside of this and has the 4th highest reputation score in the dataset. On the other side, the *@HoCPrivileges* committee has the highest reputation score of 0.43 suggesting followers of this account are the most trustworthy. This account "is a cross-party committee appointed to consider matters relating to privileges referred to it by the House of Commons"[23]. A very specific committee most likely only of interest to those who are very passionate about privileges in the House of Commons and unlikely to be targeted by bots. Therefore, while calculating the raw number of followers of an account as a way to assess effectiveness does provide some important insights it may not be sufficient in its task. Public institutions such as the UK Parliament are common targets for bots and fake profiles as a method of spreading misinformation.

An understanding of whether the Parliament's use of Twitter is effective at reaching a wide number of people can also be measured based on how many followers are shared between accounts. On average, just over 800 of the followers from any one account are also following another Parliamentary account, however this is much lower for the House of Lords accounts (red line in Figure 14). 55% of the @HouseofCommons followers also follow the @UKParliament account, while all but 6 of @CommonsHealth followers also follow @CommonsHomeAffs. In this latter case, the two accounts are almost duplicates of each other in terms of followers even though one covers the select committee in charge of scrutinising the government Department of Health and Social Care and the other scrutinises the Home Office. In this case, it could be that these two committees attract a very similar type of user, and this could suggest a latent factor combining the two which attracts the same people. Alternatively, this could be a case of poor account management by the staff running these accounts. They were both created in November 2012 and could have been advertised to the same people and at the same time, as they both also share a sizeable number of followers with @HouseofCommons and @UKParliament accounts.

This section demonstrates that while the UK Parliament uses Twitter extensively to provide information to the public, many of the followers they reach through their Twitter accounts are heavily concentrated into only a few accounts, namely the @UKParliament account. The overall reach of their Twitter accounts also does not encompass much of the UK population, however as many select committees prefer to hear from a specific set of citizens with lived experience of an issue the value of the followers lies in who they are rather than how many they are.

[23] https://twitter.com/HoCPrivileges

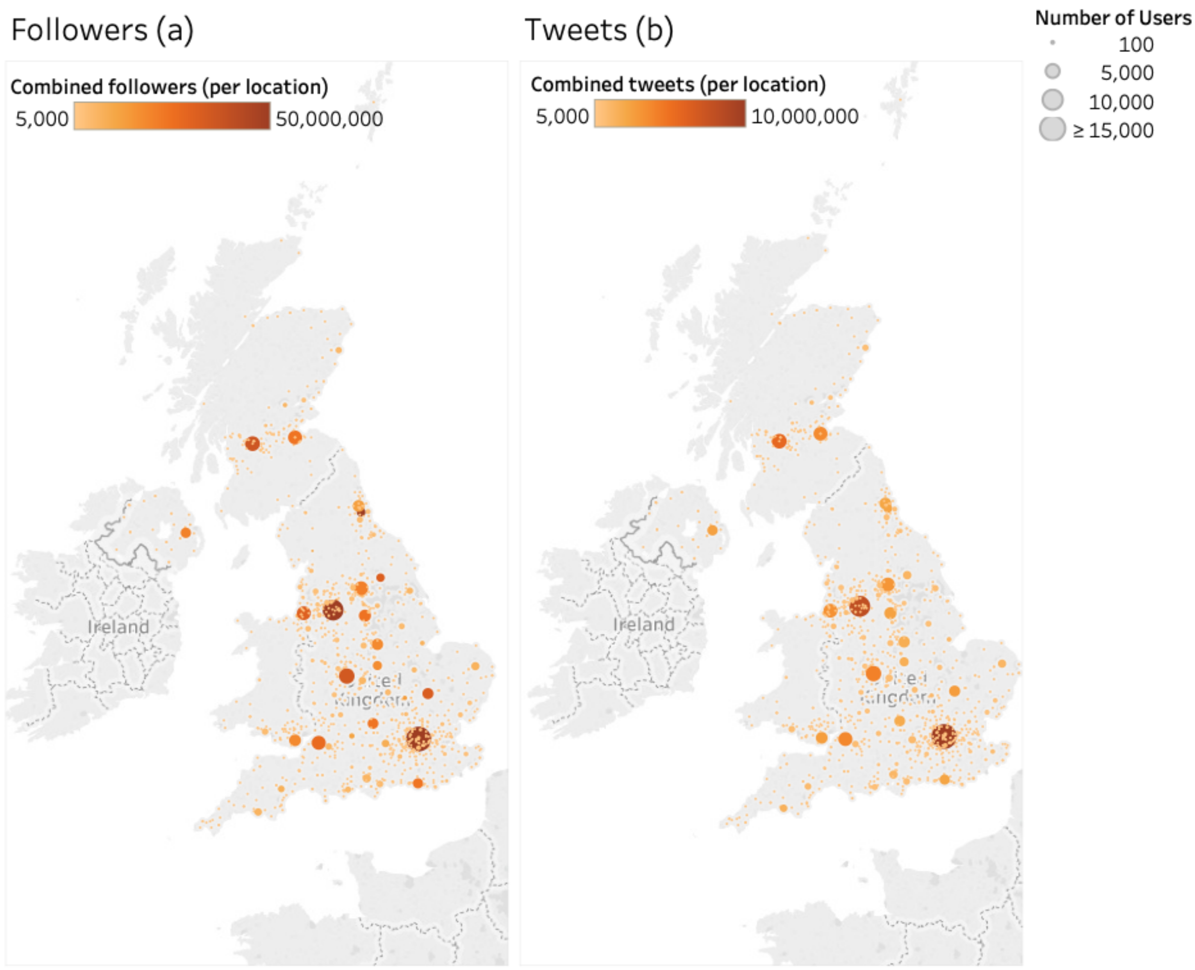
Figure 15: Average follower reputation score per Parliamentary Twitter account

5.2 Geographical spread of engagers

Understanding how many people an account is able to reach through the number of Twitter users is useful for activities such as launching an inquiry report or sharing a video. However, it is important to make use of all the features available in the online tool of choice, especially those features which can lead to valuable insights into your data. The use of social media, specifically Twitter also provides geographical information which is used to estimate the location of the tweet or Twitter user, however this is not available for the free Streaming API. This can be useful in various situations such as coordinating disaster relief efforts, for example during earthquakes or wildfires (Verma *et al.*, 2011; Sen, Rudra and Ghosh, 2015). Due to the self-reporting nature of this data, it should be used as an approximate method of analysis as the veracity is difficult to confirm. For example, only 461,287 (60%) out of 773,298 unique Twitter followers of Parliament provided their location in the dataset obtained by Parliament (described in section 3.1).

Figure 16 and Figure 17 map the location of Twitter followers, which display circles ranging in size depending on how many users are listed in a specific location, in this case at the city or country level. I can see that within the UK, the vast majority appear to be in London, but with Edinburgh, Glasgow, Birmingham, Bristol and Manchester also containing large clusters of users which is expected due to the higher populations of these cities. The colour gradient changes depending on how many followers the Twitter users in each location have. In Figure 16(a), users in London and Manchester have the most followers, however although there are not many users in York or Cambridge, cumulatively they have a high number of followers. This finding goes against the assumption that the largest cities will have the most followers. Figure 16(b) gives the same visualisation but this time for the cumulative number of tweets posted by users in each location. There are fewer unexpected patterns here than with the followers map, with London still having

Figure 16: Geographic distribution of twitter users by total followers (a) and tweets (b)

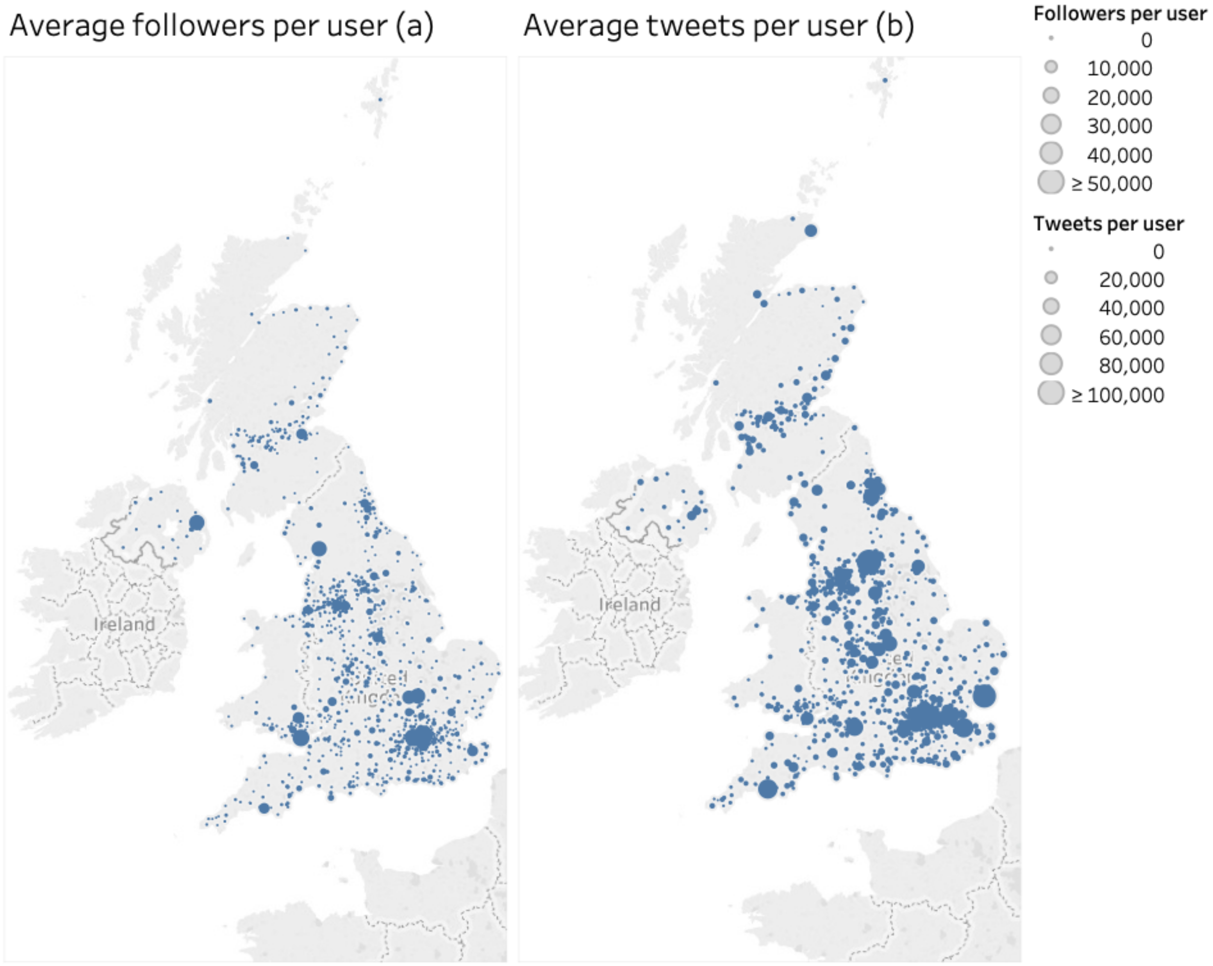


the highest number of tweets, while users in Manchester and Glasgow are also identified as big tweeters.

Although the largest proportion of users appear to reside in the UK, there are also clusters of users elsewhere in Europe. France and Spain have the largest concentrations of non-UK Twitter users. This suggests that the reach of Parliament's accounts goes beyond its jurisdiction in the UK but could be due to British emigrants living abroad.

When plotting the total number of tweets and followers, despite the exception of York and Cambridge, inevitably the areas with the largest number of people will be more prominent. To make this a more interpretative calculation, I can show the areas with the highest proportion of followers to users (Figure 17(a)), or tweets to users (Figure 17(b)) which makes the calculation more relative to how many users are in that location to begin with. This time, the sizes of the points are relative to the average number of tweets or followers per user in each location. This now shows that as well as the high-population city of London; Wick, Frinton-on-Sea, Plymouth, and Horsforth have the highest tweets-to-user ratio, while Kendal, Cardiff, and Carrickfergus have the highest follower-to-user ratio. This suggests that it is in fact the smaller towns and cities in the UK that have the most active Twitter users, especially in terms of follower numbers, while users in the larger cities are actually relatively inactive. This means although there are more users in Birmingham than in Wick, the Wick users have a much higher average number of followers themselves than the Birmingham users. These insights contribute to a better understanding of exactly who the tweets are reaching and how users in different cities behave online. Users from larger cities in this dataset are more numerous but arguably less active and involved in daily interactions making them a potentially unreliable audience for consultation, whereas the users from smaller cities may well be more inclined to participate in a more involved way (even though there are fewer of them). In recent years, Parliament has focussed on reaching citizens who live outside of London in different aspects of their engagement activities. By analysing both the raw numbers of users in each location and how those users behave on Twitter in terms of their links to other users, parliamentarians can have another measure of evaluation of their information disseminating activities. Pinpointing these locations to parliamentary constituencies would maximise the value of these insights to the parliamentary context, and allow MPs to examine how their followers interact with others online. However, to make any direct conclusions at this stage would be unwise due to the self-reporting nature of Twitter user locations.

Figure 17: Geographic distribution of twitter users by average followers (a) and tweets (b)



5.3 Effects of subject matter on engagement

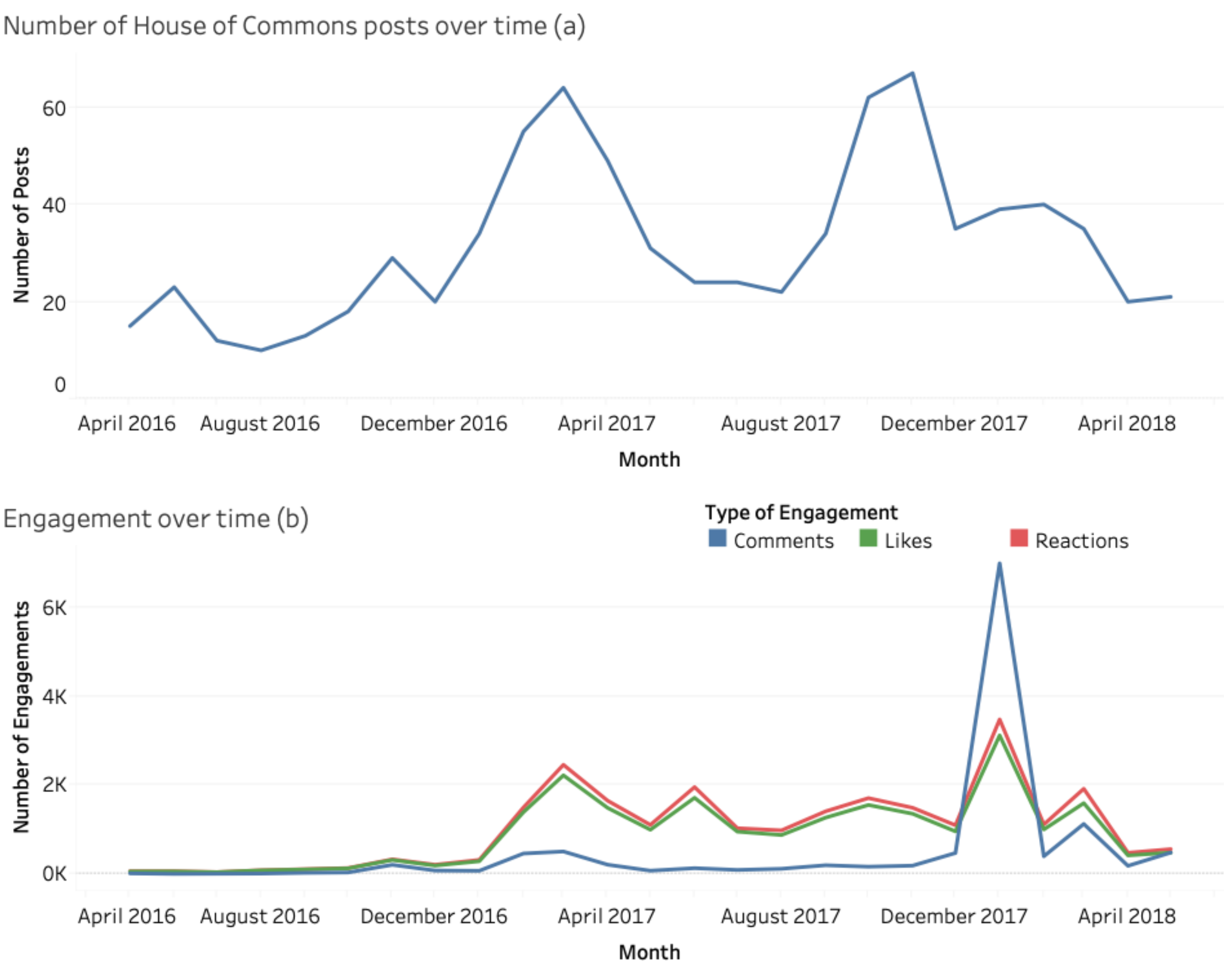
So far, this chapter has used descriptive methods to explore the Twitter data of the UK Parliament. As shown in Table 2, the Digital Engagement team use this channel of engagement primarily for activities lying in the information dimension of engagement, and the remainder of this chapter will focus on the UK House of Commons Facebook account run by the Digital Outreach team. Although this account's its primary use is to get responses and opinions from the public (more in section 6.3), it is also used partly to provide information to Facebook users. This section will evaluate the posts made by the Digital Engagement team on Facebook for the purpose of informing and consulting the public. This information dissemination is achieved through digital debates in the form of events and discussions cards in the form of photos.

5.3.1 Analysis of content of Facebook posts

I extracted all 796 posts made by the UK House of Commons Facebook account between its creation in May 2016 to May 2018. This timeframe starts just before the Brexit vote in June 2016, includes the run-up to the June 2017 general election, and continues for just under a year into the new 2017-2019 Parliament. This account is run by the Digital Engagement team who use it both to share information and collect public opinion, allowing the use of varied measures of analysis and effectiveness.

There were two major peaks in the number of posts uploaded: one in February to April 2017 and another in October to December 2017 (see Figure 18(a)). This coincides with the lead up to the June 2017 General Election, and then information about the new Parliamentary session following summer recess. However, the public's engagement with the posts (defined by number of comments, likes, shares, and reactions[24] to posts) had a definite spike in January 2018 (Figure 18(b)). This spike is most likely in relation to a debate about fireworks regulation, which took place on 19th January 2018 and received over 6000 comments. This discussion was in response to an e-petition on the same topic and is examined in more detail in Chapter 6. There are also another

Figure 18: UK House of Commons Facebook: Uploads (a) and engagement (b) over time



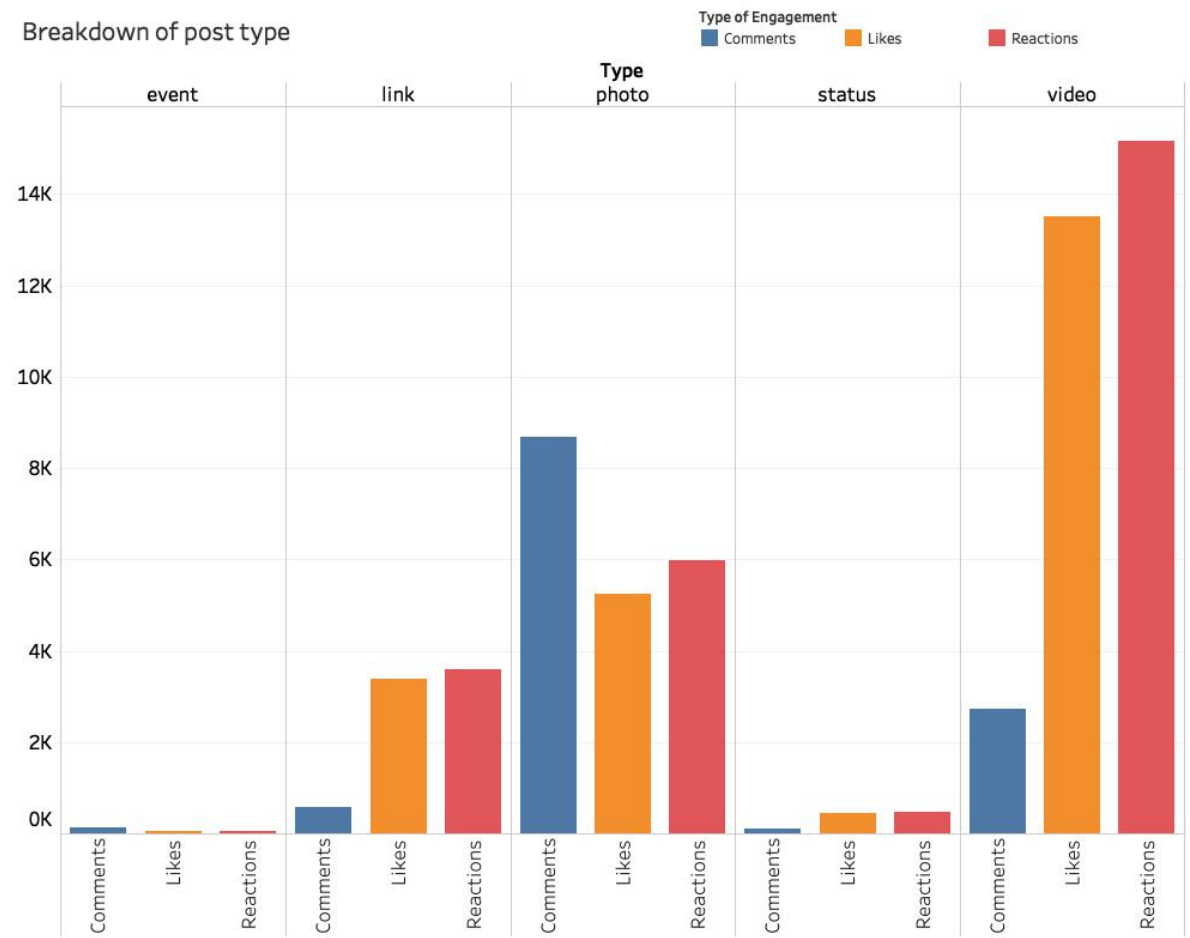
[24] Reactions are alternatives to Likes and can be "love", "haha", "sad", "wow", or "angry"

three smaller spikes of likes and reactions in March 2017, June 2017, and March 2018. The June 2017 spike is most likely in relation to the 2017 General Election, while March 2017 and 2018 could be an increased number of digital discussion cards eliciting more engagement from the public.

Of the top ten most-commented posts in this dataset: six are photos, two are videos, one is a link, and one is an event (Figure 20(a)). These posts range from topics including regulating fireworks, electric dog shock-collars, ADHD, dangerous driving, issues specific to Scotland, and the EU Referendum. The majority of these posts are to do with Westminster Hall debates led by various MPs seeking the public's views on matters important to them or their constituents. Several of the posts were in relation to e-Petitions, which had gained over 100,000 signatures and were therefore due to be debated in Westminster Hall. The one video, which received many comments was about Prime Minister Theresa May's decision to expel the Russian diplomats following the Salisbury nerve agent incident. In light of this, instead of categorising by topic, these posts can be categorised by their function i.e. the type of debate.

Where the most commented-on posts were generally related to different topics such as ADHD or regulating fireworks and asking the public for their views, the posts receiving the most number of likes are more varied. This differs from the analysis of the most commented-on posts as it explains a different type of engagement on the part of the public. It is much easier and less committal to like a post rather than take the time to post a comment (Haro-de-Rosario, Sáez-Martín and del Carmen Caba-Pérez, 2018; Bonsón, Royo and Ratkai, 2017), and shows the different types of topics which encourage the public to participate at different levels. In the top 10 most-liked posts are two photos, five videos, and three links (Figure 20(b)). The two most liked photos are also in the top commented-on posts – an e-Petition relating to firework regulation, and an e-Petition relating to holding a referendum on the final Brexit deal, suggesting these were popular both for basic engagement (likes) and more committed engagement (comments). These can shed light on the specific issues which Facebook users are most drawn to. The majority of the videos are to do with key moments in Parliament, which gained a lot of press. For example, videos marking the one-year anniversary of Jo Cox MP's death, and statements made by the Prime Minister on the Grenfell Tower Fire and the Salisbury incident. The Speaker's re-election following the 2017 General Election, and the e-Petition debate on British Sign Language also gained many likes. Finally, the links which were most liked were all to do with getting to know more about Parliament through an online course, or a podcast. These were all posted in June 2017 just before the General Election and suggests that people were eager to find out more about the institution before they cast their vote. However, as mentioned previously in this thesis, the lack of representativeness of social media data and the relatively small numbers of followers makes these results ungeneralizable to the wider population.

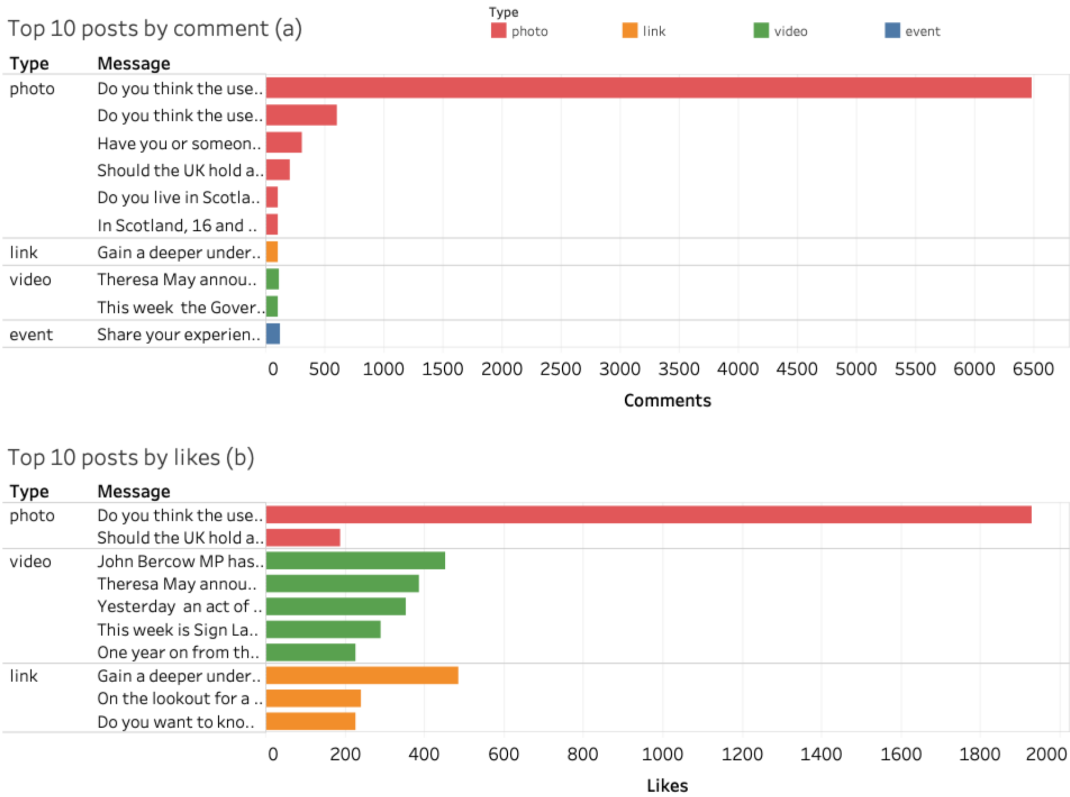
Figure 19: UK House of Commons Facebook: Breakdown of upload type



The general pattern of most-liked or commented-on posts suggests that the public are most likely to comment on posts which directly ask for their views rather than offering them up unsolicited. And Facebook users also engage well with visual media such as videos. These posts are categorised by the digital discussion cards and digital debates used by the Digital Outreach team and uploaded in the form of photos, or as an event. The public are less likely to comment on a post which directs to an external link or contains a video. As videos make up the largest proportion of posts it seems natural that they would receive the most likes and reactions of all the posts uploaded, however the data also shows that photos also receive a third of the total engagement despite them only making up 12% (99) of the total uploads to the HoC Facebook Page (Figure 19). Videos and photos were the most popular (in terms of engagement metrics) and the most uploaded types of post, while links and statuses make up a small portion of uploads and engagement (Figure 19). In other words, although only 1 in every 10 posts uploaded to the account are photos, they receive every 3rd engagement from the public.

So far, I have analysed the quantity of posts uploaded to the UK House of Commons Facebook account by the Digital Outreach team. I see that the account is actively posting an average of 32 posts a month on different aspects of parliamentary business. This particular account is used to tackle various dimensions of the engagement spectrum. Along the left branch (information dissemination), posts directing people to watch Westminster Hall debates or informing them of new online courses are extremely popular amongst the public (Figure 4). Along the right branch (encourage participation), photos and events are used to gather opinions from the public about a topic raised by specific MPs. In fact, the public are much more likely to comment on a digital discussion card than any other type of post. However, while this analysis provides a

Figure 20: UK House of Commons Facebook: Top 10 posts by comments (a) and likes (b)



useful insight into how the public react to the work of the Digital Engagement team's activity on Facebook, natural language processing can build on this preliminary exploration to ascertain exactly what is being said and how the public is reacting to the content they are being provided.

5.3.2 Delving deeper: Natural Language Processing

Moving onto the content of the Facebook posts made by the UK Parliament account, there are various ways to summarise large volumes of textual data. One of these is through the analysis of the most frequent individual words (unigrams) or pairs of words (bigrams), which are displayed in Figure 21. This shows that the most common bigram (consecutive word pair) used is a link to the parliament.uk webpage (bit.ly) showing there is a strong link between the Facebook page and the Parliament website. Westminster Hall and Prime Minister are also common bigrams. Figure 21 shows that there is a large cluster relating to different departmental select committees showing that a wide range of committee news is covered by this account. There is also a small cluster relating to the watching of MPs debating in Westminster Hall and to the links to the parliamentlive.tv website (where parliamentary sessions are live-streamed). This network provides a glimpse into the content of the posts uploaded by the Digital Engagement team and reveals a wide range of themes within the posts on the account. Unlike accounts owned by Members of Parliament, this account must be politically impartial and represent the views of different parties and MPs including all parliamentary groups.

By mapping the network of bigrams in this way, I can understand if that task of impartiality is achieved and ensure that the public is getting an impartial and comprehensive view of parliamentary business. Evidence of the timeframe of the data can also be seen in the network with links such as "Northern Ireland", "European union", "Withdrawal Bill" and "June 8th" which all relate to the Brexit vote. "Online discussion rules" relates to the digital discussion cards and Digital Debates held as photos and events (examined above), while "people signed" is in reference to the e-petitions the account frequently uses as a basis for discussion cards or to link to Westminster Hall debates.

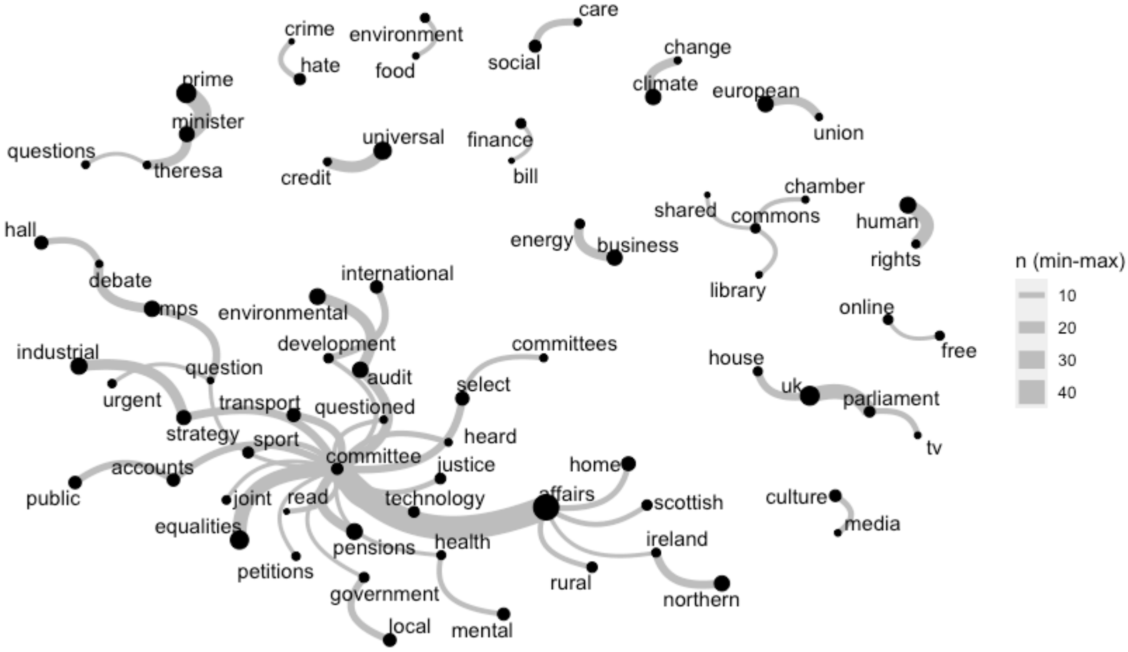


Figure 21: Bigram network of HoC Facebook Page

Whereas the analysis of bigrams shows an overview of the posts based on word pairs which are used most frequently, Figure 22 shows the top 5 words associated with 18 topics extracted by an LDA topic model, and Table 4 suggests the prevalence of each of the 18 extracted topics and their interactions. This LDA model has a low alpha value of 0.0187 and suggests the posts are represented by just a few topics and are quite different to each other. The interactions are identified as the number of times a post has been liked, commented on, or reacted to by Facebook users. The number of topics was chosen through evaluation of a range of models (Nikita, 2016) (further details in section 3.2.3). The topic model results show that the reach in terms of subject matter and content covered by the UK House of Commons Facebook account touches many aspects of parliamentary business, as is the intention of the page.

Figure 22: UK House of Commons Facebook posts: Overview of LDA topic model (posts made

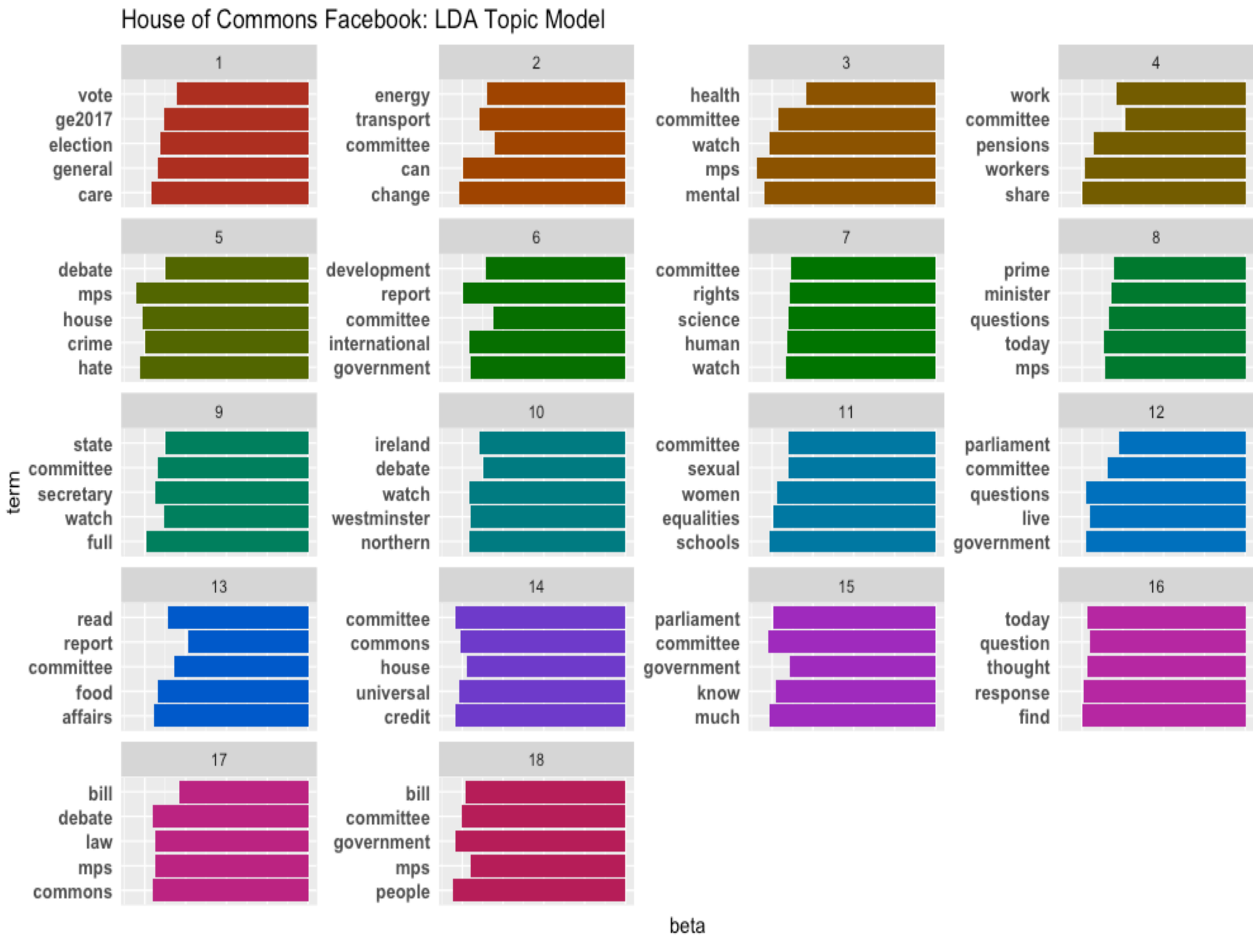


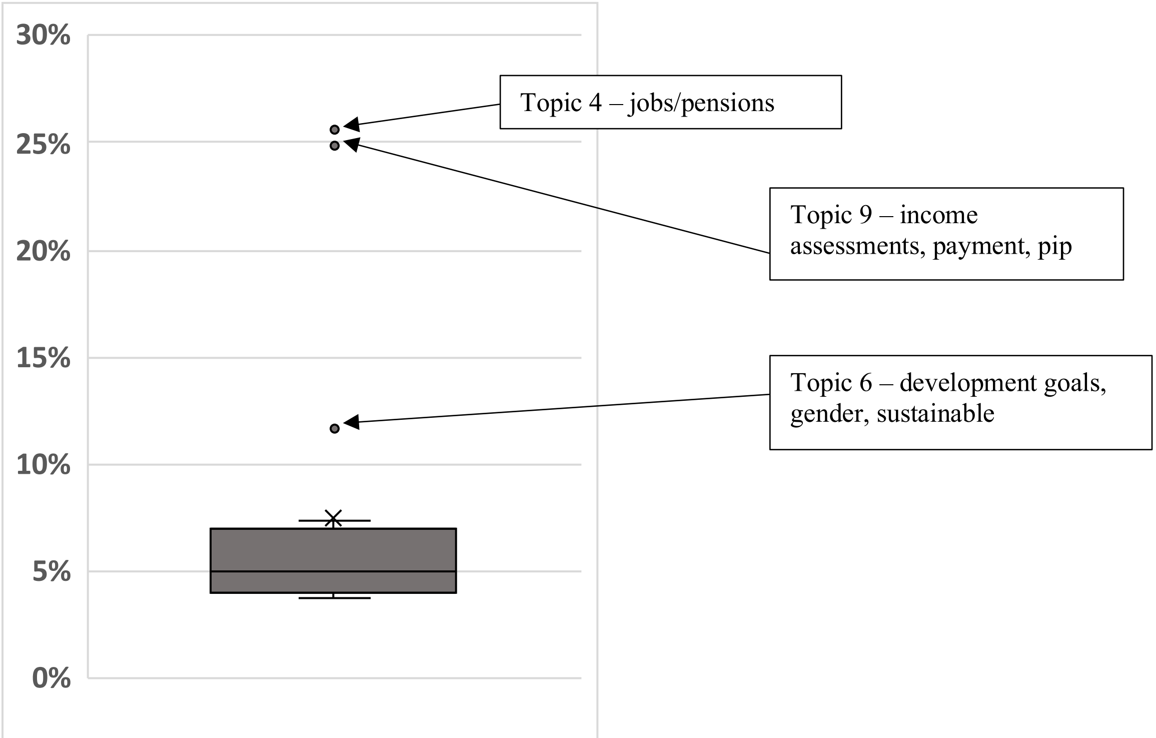
Table 4: Topic descriptions for House of Commons Facebook page

Topic Number	**Main terms/themes**	**Percentage of posts**	**Percentage of interactions**
1	Voting, GE2017	9%	13%
2	Energy, climate, industrial, strategy	7%	3%
3	Health, mental	4%	1%
4	Jobs, pensions	4%	2%
5	Hate crime,	4%	6%
6	Development goals, gender, sustainable	4%	3%
7	Human rights, science, data, referendum	6%	4%

8	PMQs	10%	8%
9	Income assessments, payment, pip	4%	5%
10	Northern Ireland, Batley & Spen constituency, republic	4%	4%
11	Sexual harassment, teachers, equalities cttee, royal marines	4%	2%
12	Petition, online course	5%	9%
13	Environment, food, budget, Hammond	6%	5%
14	Modern slavery, universal credit, victims, asylum	4%	7%
15	miscellaneous	8%	7%
16	Rail travel, PMQs	4%	5%
17	Loneliness, Brexit, finance, withdrawal	6%	4%
18	Employment, broadband, Scotland	6%	4%

Looking closer at the topics from an LDA model displayed in Figure 22 I can analyse the textual content of the data. Table 4 shows that topics 1 and 12 were the most popular with users, having over 1500 individual users interacting with each. These two topics relate to voting in the 2017 general election and online petitions and courses respectively. Posts relating to Prime Minister's questions were also popular having just under 1500 users interacting with them (topic 8). While these topics received the most likes, comments or reactions from users across the two years, it does not necessarily mean that these topics attract the most engaged users. The network data not only captures whether an interaction took place, but also the number of times users interacted with a post. For example, one user could like and comment on the same post, or respond to another user several times on the same post. Figure 23 shows the distribution of difference between the number of interactions with a topic and the number of users engaging with that topic. For example, if for each topic a user engages with, they only interact with it once (e.g. 1 like), the percentage difference is 0%. Any further interactions with the post by the same user would increase this percentage. This is shown in the equation, $\text{Average interaction} = \frac{\boldsymbol{Interactions - Users}}{\boldsymbol{Users}} \times 100$, and calculated for each topic.

Figure 23: Proportion of Facebook post interactions to Facebook users by topic defined by LDA



The boxplot shows that the majority of the topics have between 4% and 7% difference between interactions and number of users involved, meaning users interacted on average 6% more with each topic than just liking or commenting on the post once. Exceptions to this are topics 4, 9 and 6 which had over 26%, 25%, and 12% difference respectively. These topics give an insight into the types of issues that encourage the most engagement. Despite these topics being the most engaged with relative to the overall popularity, they are not the most popular topics in the whole corpus. This analysis of user interactions with topics show that while users may find one particular topic or issues most interesting overall (in this case voting and the 2017 general election), it is actually the less popular topics which attract the most involved users in terms of how they interact with a post. In the UK House of Commons context, I see that financial issues relating to jobs, pensions and income assessment receive proportionally many more comments or likes compared with other topics. Issues surrounding the environment also fall into this category. The Digital Engagement team can use these insights to asses which of the topics they provide information about attract the most enthusiastic users and therefore could encourage the best online discussions for MPs and future engagement sessions.

5.3.3 Delving deeper: Network analysis & community detection

Analysing the reach of information dissemination and textual analysis of the posts by the UK House of Commons Facebook account provides valuable insights into the functioning of the account and how the public engages with it. However, the analysis completed so far does not give any information into how the public engage with each other through this account. Questions such as "are there certain posts or topics which attract the same people?", "do I have any trolls engaging with the account", or "who are the most influential people in the account?" require a combination of textual and social network analysis to answer.

Using the Netvizz Facebook application, I was able to extract the social network data for the House of Commons Facebook page from 17 May 2016 to 15 December 2017[25] . The data allows me to create a bipartite network graph of 8,343 nodes, 7,717 (92%) of which were users and 626 were made up of different posts – either photo, video, link, status, or event. These can be linked through 18,623 edges which recorded any interaction made between a user (source) and a post (target). These edges also have a weighting which represents the number of times a particular user interacted with a specific post. There were 6 posts that received no interaction from the public. These were:

1	*How effective is the British Transport Police in tackling crime on the railway? The Transport Committee are holding a session in its inquiry into rail safety, and are hearing from the British Transport Police today at 4.05pm. Find out more information and watch live: https://goo.gl/QCuAWl*	16/01/2017
2	*UK House of Commons shared Arthritis Research UK s event.*	19/10/2016
3	*Watch the Prime Minister, David Cameron, make a statement and answer questions from MPs in the House of Commons following the results of the EU referendum.*	27/06/2016
4	*How can the rail passenger experience be improved? Watch the Transport Committee question experts as part of their inquiry into improving the passenger experience: http://goo.gl/upqFAm*	06/06/2016
5	*The Environment, Food and Rural Affairs Committee is holding an inquiry into the welfare of domestic pets. Watch them question experts on the sale of pets online.*	25/05/2016
6	*Day 4 of the Queen's Speech debate focused on Europe, Human Rights and keeping people safe at home and abroad. The Counter-Extremism and Safeguarding Bill will aim to provide stronger powers to disrupt extremists and protect the public. Watch what MPs said about the proposed legislation.*	24/05/2016

These posts range in subject matter and the date they were posted (from May 2016 to January 2017), however they share a characteristic that encourages users to watch a video as a link rather than an embed. This supports the interpretation of Figure 19 which showed that links received much less engagement than videos or photos, and suggests the public's choice whether to engage or not with a post could be more related to the way the post was created rather than the subject matter. It is much simpler to view an embedded video when scrolling the Facebook homepage than

[25] Facebook restricted the Netvizz application in early 2018 making it impossible to extract any further data.

to click on a link which redirects the user to another webpage. A link creates an extra step for the user to take to view the content and as such increases the barrier to engagement.

As explained in section 3.4, several measures can be applied to networks, each giving a different interpretation of the data and revealing different patterns in the network. Degree centrality measures which nodes have the most connections or edges in the network. The density calculates the proportion of connections in the network out of all possible connections that could be made (Dean, 2018).

Focussing firstly on the posts projection in Figure 24, I can visualise the nodes based on betweenness centrality (node size) and Louvain modularity class (node colour). Three different community discovery algorithms were used to partition the data: Girvan-Newman, Louvain (GN), and Label Propagation (LP) (Bedi and Sharma, 2016; Newman and Girvan, 2004) (see Chapter 3). Figure 26 shows the number of posts allocated to each community for the three algorithms. It shows that GN and LP algorithms categorize the vast majority of the posts into a single community (GN1 and LP1 respectively). This suggests that the posts broadly are about one single topic. Louvain on the other hand categorizes posts more evenly between five of its eight communities suggesting more of a uniform distribution which is more reflective of the bigram network seen in Figure 21 and the optimum number of 18 topics seen in Figure 22. The bigram network shows a range of issues being raised in the posts made by the House of Commons and therefore suggests a topic model should also recognise this diversity in topics.

Looking back to Figure 24, one node (Post281) has a relatively low degree of 47 but the 4th highest betweenness centrality, which suggests that while this post 281 did not have many interactions, the ones it did have were very central to the overall popularity of the post. This post was entitled:

> "*Gain a deeper understanding of how the UK Parliament works including Bills, Select Committees and the roles of the Speaker and Lord Speaker by taking our FREE online course. Starts 27 February. Sign Up Now! goo.gl/d4QaxI*"

Posts 410 and 025 had the highest betweenness centrality suggesting they intersected many other nodes. These posts related to a petition to reduce tuition fees:

> "*Should tuition fees in England return to ¬£3 000 a year? More than 164 000 people signed a petition saying that they should. MPs are now debating it in Westminster Hall. Watch the whole debate here: goo.gl/8h4wjb"*,

and a post about reducing the House of Lords membership:

> "*Does the House of Lords need to change? With over 800 members it's the largest second chamber in the world. The Public Administration and Constitutional Affairs Committee is asking how the House of Lords size and make-up can be managed. Find out more: https://goo.gl/3ggjMR*".

The Louvain community detection algorithm has a relatively low modularity of 0.3 giving 5 distinct clusters of posts determined by the number of connections nodes have to each other. This suggests there are distinct communities of posts which have been interacted with by the same user suggesting at least five distinct clusters of topics which users are interacting with. By examining the topic numbers from the LDA topic modelling which are most prevalent in the clusters identified by the Louvain modularity algorithm, I see that the green and grey clusters are dominated by topic 1 (voting) and topic 8 (PMQs) respectively. These two clusters are adjacent to each other in Figure 24 and suggests users interacted with both of these topics. The purple cluster is dominated by topics 18 (employment) and 2 (climate/energy), while the blue cluster mainly contains posts in topic 17 (loneliness and Brexit). The orange cluster does not have a dominant topic in the posts but is instead comprised of a range of different topics and issues raised by the UK House of Commons Facebook account. So, the different areas of interest of users in this network show a subset of people interested in core parliamentary business such as PMQs and voting in the general election, and another subset of people more concerned with issue-related topics such as loneliness, and employment which are not exclusively related to the UK Parliament.

Figure 24: Facebook post network nodes weighted by betweenness centrality and clustered by Louvain algorithm

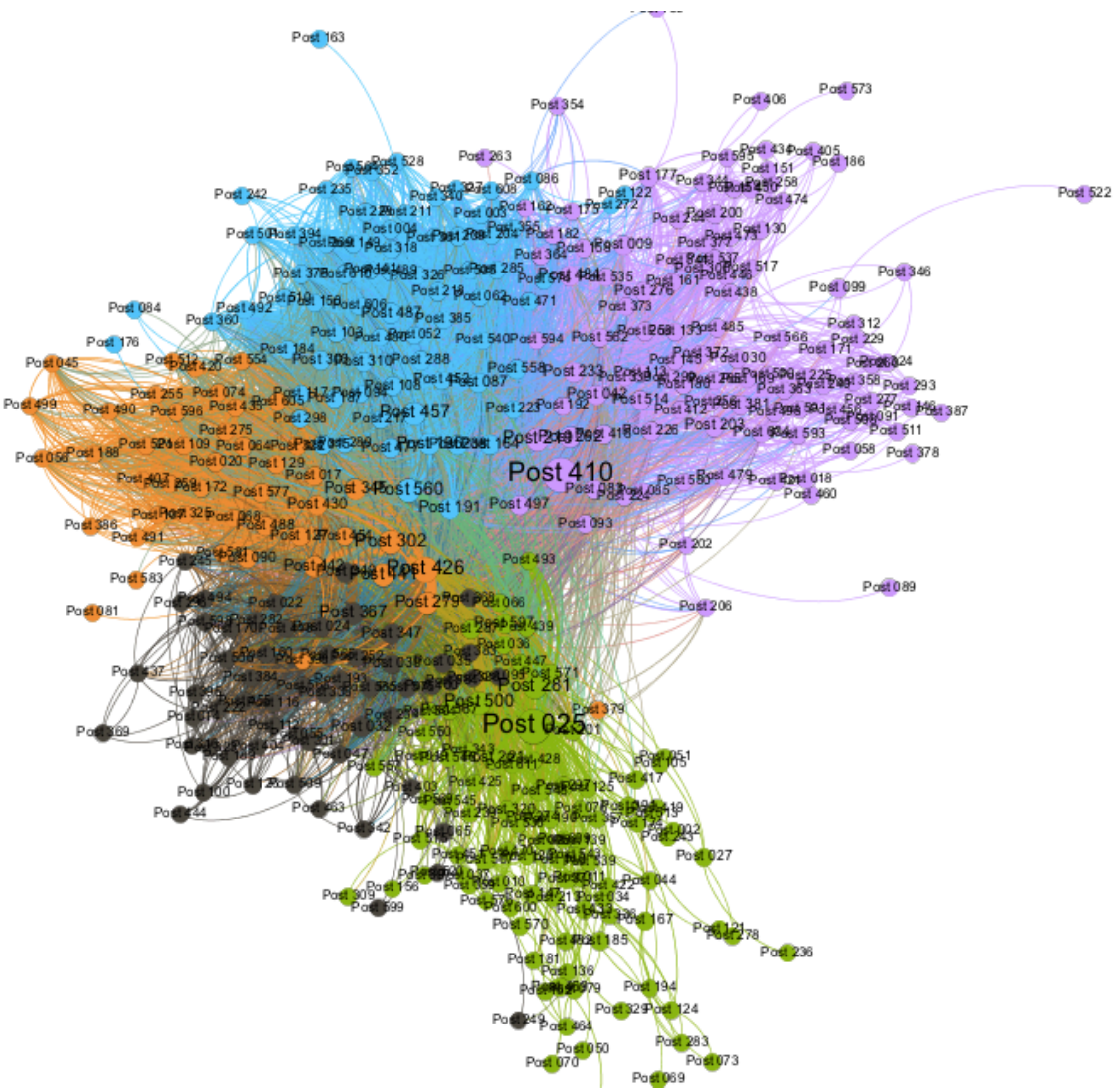


Figure 26: Community detection algorithm results of Facebook post distribution for Girvan Newman, Louvain method, and Label Propogation algorithsm

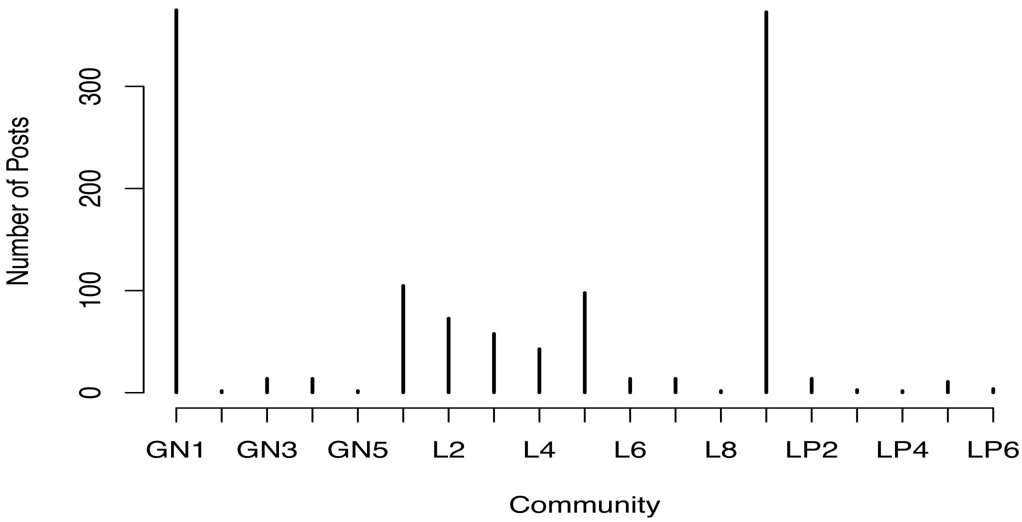
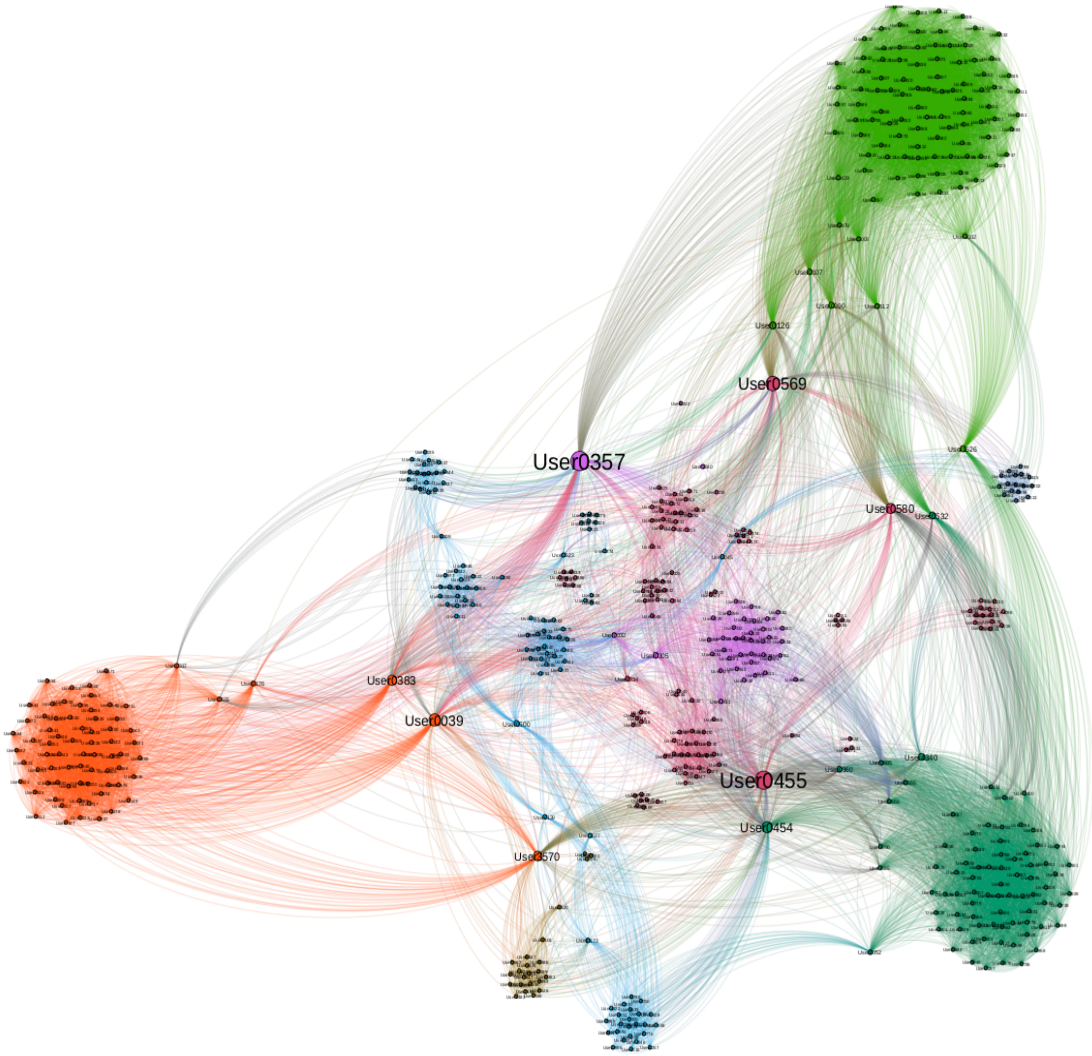


Figure 25: Facebook user network: nodes weighted by betweenness centrality and clustered by Louvain method



Focussing now on the user projection, Figure 25 displays a graph with nodes representing

Facebook users and edges representing users who have interacted with the same topic. This was achieved by aggregating the posts into one of the 18 topics found in the LDA topic model. This graph has a Louvain modularity of 0.64 meaning the communities were relatively well defined and an average degree of 60. Figure 25 shows that there are three large clusters of users (orange, light green, and dark green) with several smaller clusters connecting them. Most users can be grouped alongside others who are interested in the same topics as them, while some users have varied interests which do not fit into one single cluster but bridge several ones (e.g. User0455, or User0357).
Social network analysis allows for an exploration into how nodes interact with each other and how they form clusters based on these interactions.

The pattern of clustering in Figure 25 shows that most people share interests in various topics as to be expected, for example, people who are interested in certain topics are also likely to be interested in other specific topics. The smaller clusters of users also suggest that there are small interest groups which cover a smaller range of topics and therefore attract a small set of users.

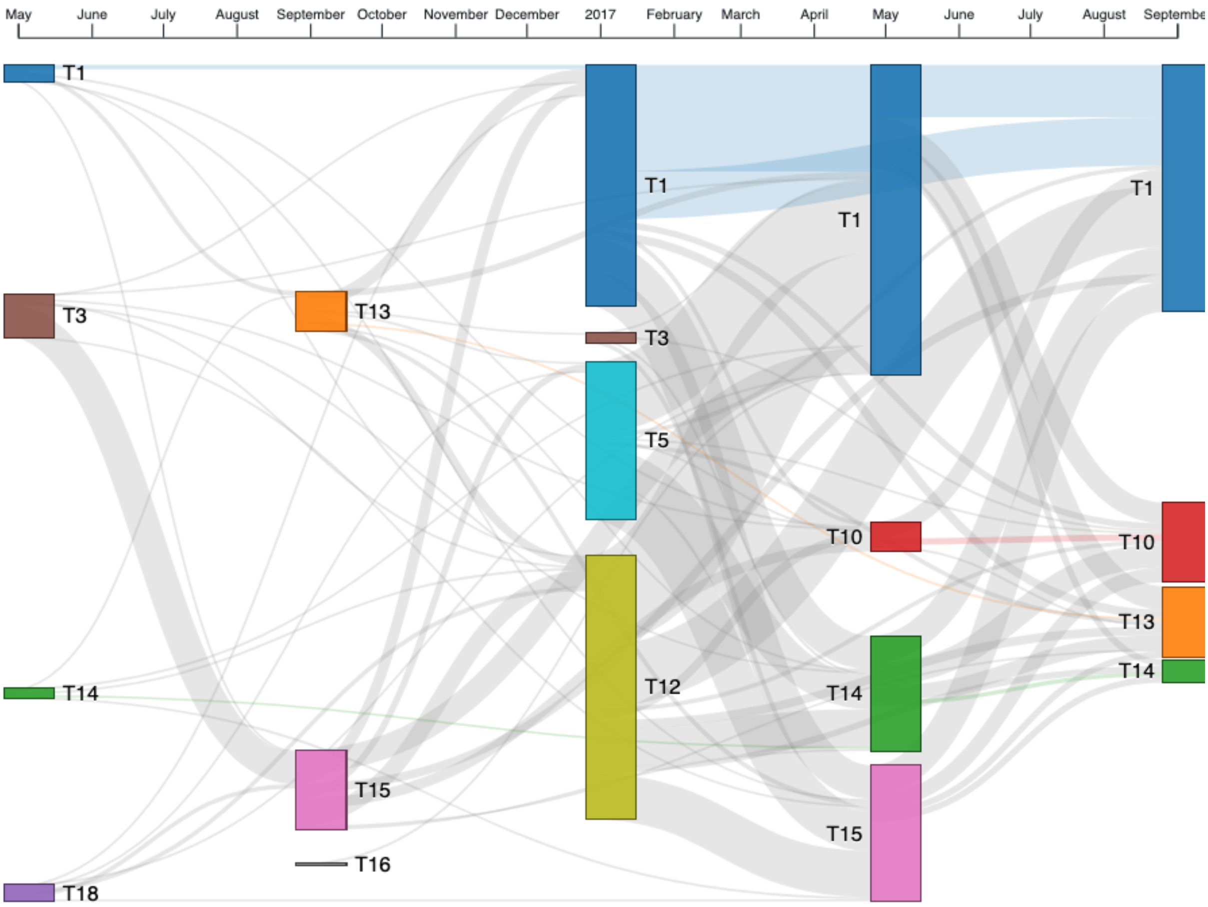
The large nodes User0357 and User0455 in Figure 25 have high betweenness centrality values meaning they intersect many edges and could act as 'influencers' in the network. Nodes with positions in the network such as this are referred to as cut-points. These are defined as the nodes which bridge different community clusters, and if removed would cause the network to split into further smaller clusters. These are therefore seen as the nodes which hold the network together and connect its different communities. User0357 is especially interested in many different topic areas and its high betweenness centrality suggests they are very active in liking or commenting on many different posts. Users such as this one may be interpreted as trolls if the interactions they are making are particularly disruptive. As well as individual users acting as cut-points, there are small groups of users that can act as cut-points too (e.g. purple or blue clusters). For example, the right-most dark blue cluster (at 3 o'clock) is a small community of Facebook users that may have a particular niche interest in a specific topic. This topic does not attract any other users, however the people within this community are also engaging with the same posts/topics as those users in the larger light green and dark green clusters. This suggests that while the majority of users are interested in similar topics, there are some smaller groups of users whose interests span different topic areas.

While this analysis gives an insight to the network of Facebook interactions between users an UK House of Commons posts at a specific snapshot in time (December 2017), the network data also has a longitudinal feature. As mentioned at the beginning of this section, the data spans several important events in the British political calendar, the 2016 Brexit referendum and the 2017 general election. Creating a dynamic network taking into account the date of interaction would allow me to see if the pattern of interactions based on subject topics changed due to one of these political events. For example, the large orange cluster in Figure 25 could have previously been made up of several smaller topic clusters which have combined over time. Conversely, the green and turquoise clusters could have been one large cluster before the General Election which have now separated.

Following pre-processing steps of the network data as described in section 3.4.2 I am left with 542 remaining user nodes who had interacted with multiple posts over multiple time periods. This alone shows that although many users interacted with individual posts over the two years, only a small proportion of them were continuous engagers over time. I found the best time frame was to use a quadrimestre (4 months each), creating five time periods (Figure 27). I found that those users interested in voting and Brexit (T1) between May-August2016 did not interact with any similar voting posts leading up to the General Election in January-April 2017. This suggests there are two distinct groups of users on Facebook, one interested in posts to do with Brexit, and another to do with the General Election, and is contrary to the general idea that users interested in Brexit would also have an interest in the Election. For the year 2017, many users who were originally interacting with posts to do with voting and the election remained interested in these topics up until December of that year. This could reinforce the value of information dissemination, specifically education type of initiatives in leading to more engagement.

Other topics which saw a continuation of interest between one time period and another was to do with universal credit (T14), and Northern Ireland/Batley and Spen (T10), the latter most likely related to Brexit vote and negotiations. The majority of other topics received interest from users who had previously been interested in other topics. This suggests that aside from those heavily interested in voting, most regular visitors to the House of Commons Facebook page were interacting with a wide range of different topics, likely more based on issues close to them. However, the vast majority of users still only interacted with the Facebook page once without repeated likes or comments on another post in the quadrimestre time periods.

Figure 27: Sankey chart of Facebook user interaction with UK House of Commons Facebook page topics (T) between May 2016 and December 2017



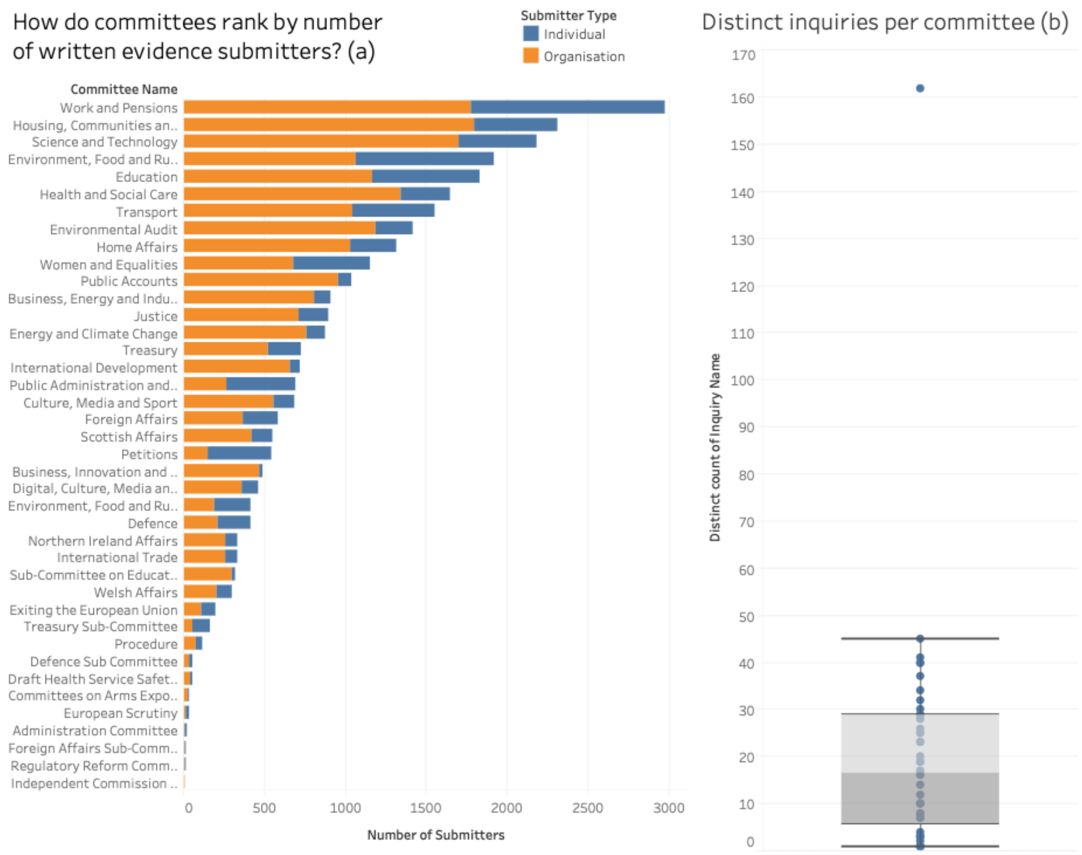
This shows while using a quantitative analysis of number of comments is necessary, it is not sufficient for examining the true levels of engagement with a particular post or topic. Other evaluation measures such and exploration of topics and dynamic social network analysis provides a deeper understanding into how the posts made by the Digital Engagement team to provide information to the public and also solicit views can be assessed.

5.3.4 Select Committee evidence networks

As well as assessing the data from online engagement sessions, select committees are a valuable source of data from formal engagement activities. They have a unique position in Parliament where they are able to bring together cross-party MPs to scrutinise the government through their inquiries into specific issues. As part of this, they frequently contact the public through calls for evidence which are statements of support for the public's views on the inquiry topic. These are usually submitted in the form of word document of a maximum of 3000 words (Parliament.Uk, 2020). Section 2.6, outlined some of the work done by select committees and highlighted some of the methods used to innovate the way they receive evidence from the public. However, while these new methods are being introduced, the public still have the option to submit formal evidence to different inquiries through a written document. Along with managing many of the engagement activities for various select committees, the Web and Publication Unit (WPU) has a goal to produce macro-analyses of committee evidence submissions to inform stakeholder engagement strategies. Through contacts in the Select Committee Office and WPU, I was able to obtain anonymised data from written evidence submissions between September 2016 and September 2018. The purpose is to identify how evidence submitters move (or do not move) between inquiries and committees, and examine the diversity of those submitting evidence.

In total over 30 thousand submissions were made, 21,549 (71%) by organisations, and 8,686 (29%) by individuals to 19 committees. The majority of committees conducted between 5 and 45 inquires each with an average of 43 submissions in each inquiry. As Figure 28 shows, the majority of committees have more submissions from organisations than from individuals, with the exception of the Petitions Committee, Public Administration Committee and the

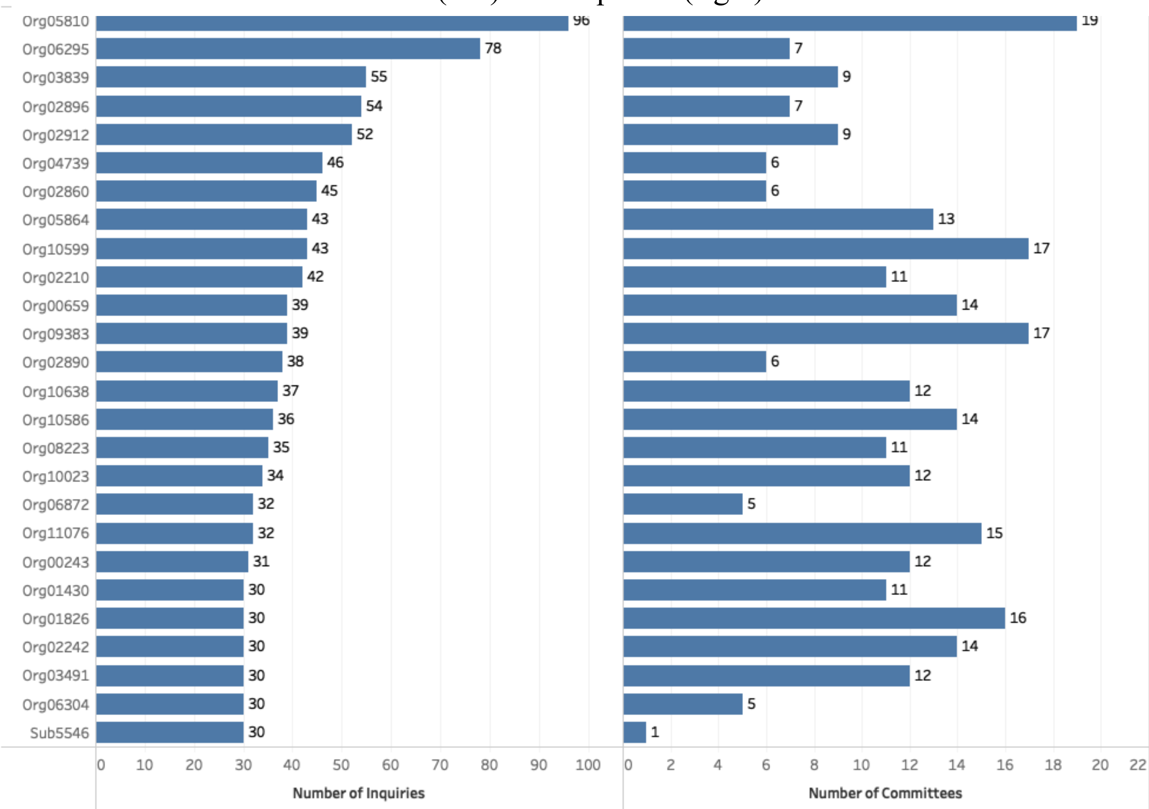
Figure 28: Select Committees ranked by number of written evidence submitters (a) and number of inquiries per committee (b)



European Scrutiny Committee. Work and Pensions committee by far have the highest number of submissions overall (Figure 28a), but I cannot assume a high number of inquiries will directly result in a high number of submissions to those inquiries. The Public Accounts committee was unique in this dataset as it had 162 distinct inquiries (Figure 28b) over the two years with an average of 6 submissions in each. On the other hand, the Environment, Food and Rural Affairs Sub-Committee, had an average of 139 submissions to only 3 inquiries and the Work and Pensions committee had an average of 74 submissions per their 40 inquiries. Therefore, the high number of submissions to the Work and Pensions committee seems to be because of the relatively high submission to inquiry ratio of this committee rather than it having a very high number of unique inquiries. Nevertheless, there is a great variation in the number of inquiries and submissions of each of the committees in this dataset.

Witnesses making the most number of submissions are almost exclusively categorised into organisations, with the exception of one individual (Sub5546) who made submissions to 30 inquiries, all to one single committee (Work and Pensions). Org5810 (Local Government Association) made submissions to 96 inquiries within 19 different committees (Figure 29). Another unique committee is the sub-committee on Education, Skills and the Economy which had only 2 inquiries: one on Apprenticeships receiving 184 submissions, and another on Careers education, information, advice and guidance which received 132 submissions. 95% of these submissions were made by organisations with the remaining 5% from individuals. Their high submission to inquiry ratio suggests these inquiries were especially important to the public, specifically businesses and organisations.

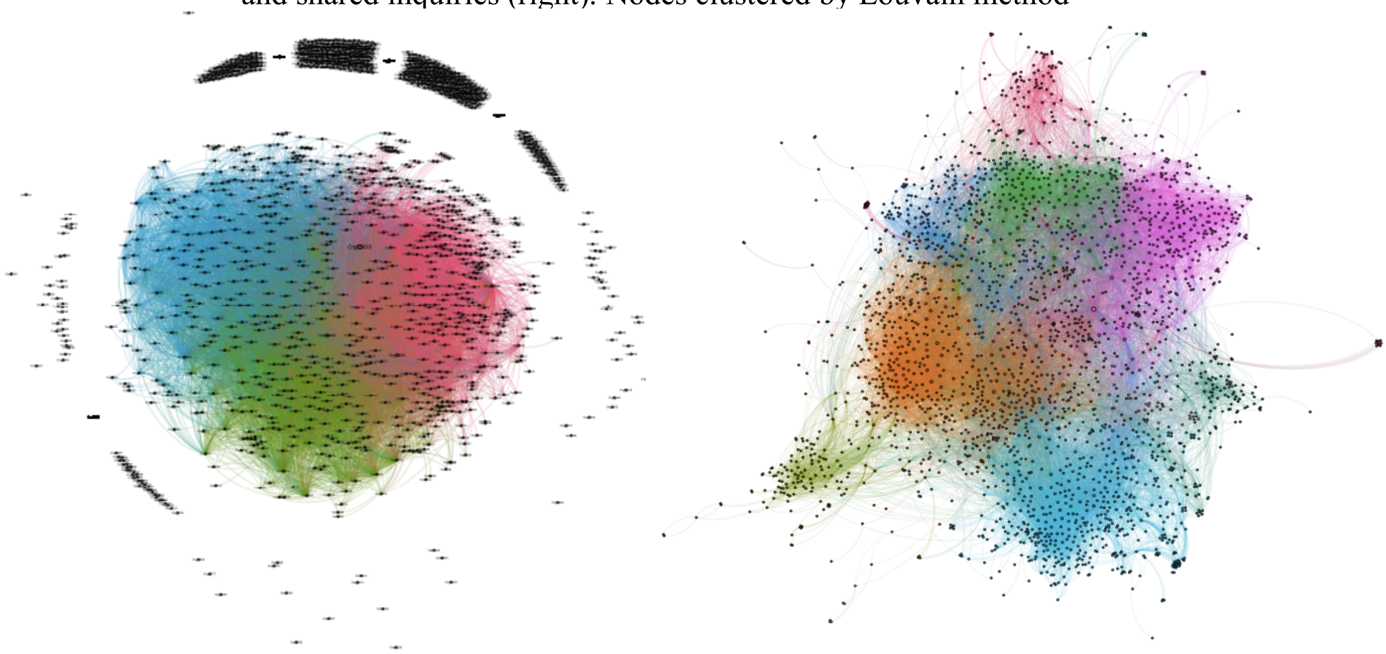
Figure 29: Written evidence submitters ranked by number of submissions to committees (left) and inquiries (right)



Through social network analysis I can also show how the submitters to inquiries group themselves based on either the committees or inquiries they submit evidence to. Figure 30 (left) shows a network graph where each node represents an individual or an organisation. These people/organisations have a connection if they have submitted evidence to the same committee. Organisation 5810 (pink cluster) is the largest node (highest degree) meaning they have submitted evidence to the same committees as many other organisations or individuals. However, the low Louvain modularity score of 0.225 shows witnesses do not cluster well based on committee structure. Taking density into account, the relative modularity of this network as defined in section 3.4.1 is 0.247.

Clusters are much more apparent and well defined when they are grouped on the basis of inquiries (Figure 30 right) rather than committees, as this network has a Louvain modularity of 0.56 (relative modularity of 0.52). Distinct groups of witnesses can be seen based on general topics of inquiries, for example, the green cluster (7 o'clock) represents people who submitted to defence topics. These networks show that people are more inclined to submit written evidence to Parliament based on the topic of the inquiry rather than adhering to the internal committee structure. This finding is supported by literature which also finds that people engage because of issues close to them rather than for any other reason (Liaison Committee, 2015; Hansard Society, 2011; OECD, 2009). The difference in modularity and relative modularity of both networks also shows that despite the different network densities, the clustering of witnesses by inquiries rather than committees is still more effective. These topic clusters of witnesses could suggest new avenues for communication or engagement, such as new social media accounts based on the topics users share an interest in, or special-issue newsletters for example. There is also an opportunity to share different inquiries with past evidence submitters who have already shown an interest in a particular issue, even though they may not have encountered a committee before. This sharing of data and resources among parliamentary teams is something I put into practice during my experiments with online discussion platforms explained in section 6.5.

Figure 30: Written evidence submitters grouped by shared committees (left) and shared inquiries (right). Nodes clustered by Louvain method



Conclusion

As has been raised in previous chapters, public engagement in the UK Parliament can be categorised into two dimensions; informing the public and encouraging participation. This chapter has focussed on the former dimension exploring where the UK Parliament's use of social media in digital engagement. I have specifically addressed the use of Facebook and Twitter as a method of providing information to the public following Parliament's accounts on those platforms

Section 5.1 examined 48 Twitter accounts owned by various teams in Parliament to show that due to the different dates the accounts were created, they differ heavily in number of followers, with *@UKParliament* having by far the most. However, I also see how a large number of followers does not always mean that those followers are legitimate in terms of their reputation score, and the characteristics of these followers must also be taken into account when possible. The geographical spread of these followers can also provide insights into how effective an activity can be at reaching a wide range of people. For example, despite Parliament having a small number of people in York following their Twitter accounts, these users have a proportionally higher number of people following them. Meaning, should someone from York like or retweet a post from the *@HouseofCommons* account, that post is likely to be seen by even more people than if it had been retweeted by another user. I can begin to use these insights to build a picture of the Twitter followers of different accounts, assessing how their profile characteristics or attributes can influence their reactions to different posts.

In section 5.3 I examined the content of posts uploaded by the UK House of Commons Facebook account. Textual analysis of these posts revealed the teams managing this account inform the public of a wide range of different issues. Analysis of the interactions between users engaging with the Facebook account show that even though the public may engage very little with a topic, exploring the interactions between the public and that topic can reveal unlikely communities of passionate participants who can make valuable contributions to a discussion. Furthermore, the network of interactions reveals there are several individual users who engage with different topics, however the majority of users interact solely with others who interact with the same topic, and do so just once rather than repeatedly throughout the year.

Geo-spatial analysis in section 5.2 found Parliament's current digital engagement activities are effective in reaching a wide range of users around the UK and providing an impartial account of different aspects of Parliamentary business, however the legitimacy of these users should be taken into account as evidenced by the analysis of Twitter user reputation scores. The Facebook analysis shows that the public respond well to this channel of engagement by grouping themselves in ways that align with the varied topics and themes posted. While this suggests a worthwhile use of the account, if the goal is to encourage more interaction between people, creating content with more crossover between topics may be more effective. Parliament is undertaking engagement at all levels, from providing information to collecting responses. However, the extent to which the more participatory online activities work to encourage a range of different people to engage and to understand their contribution requires further investigation, starting in the next chapter.

Chapter 6 Citizen Engagement: Who are they and what do they want?

This chapter focusses on analysing the participatory online engagement activities, digital debates and social networks of actors involved in some of these activities. I use text mining and social network analysis to reveal information encoded into the data. I will show how different debate topics lend themselves better to different methods of analysis and explore the reasons behind this. Where previous chapters explored the information dissemination forms of digital engagement, this chapter focusses on activities solely designed to encourage discussion and gain views from the public. These activities lie within the Encourage Participation branch of the spectrum of public engagement (Figure 4) and are primarily led by the Digital Engagement team.

Since their introduction in 2015, digital debates have been held on various social media and microblogging platforms providing a large amount of data from the public (Digital Democracy Commission, 2015). This data is in the form of textual comments posted on different topics proposed by Members of Parliament, and includes the public's opinion on these topics in preparation for a debate or inquiry. Currently, these comments are individually read and summarised by the Digital Engagement team or select committee staff to send to the Members of Parliament, who can then incorporate the viewpoints of the public into their parliamentary speeches. When a particular discussion is also covered on Twitter using a specific hashtag, the Digital Engagement team has no means to capture the relevant tweets and to analyse the network of Twitter users engaging with a given topic. As a result, vital data is not being used to its full potential and useful insights are being missed.

The Digital Engagement team usually has a range of questions to answer when analysing a digital discussion. While the answers to these questions are primarily for the Members of Parliament who will use the results in their debates, the team also uses the analysis of the discussions to inform the organisation of future engagement sessions[26]. The process of manual analysis may be sufficient for digital debate topics which only receive a limited number of comments, however there have been several discussions which have garnered hundreds or thousands of comments. Reading each comment separately creates a large drain on staff resources, especially as the turn-around time for summary reports is often very short due to the fast-paced business of Parliament and lack of time between the online discussion and chamber debates. In some cases, this has caused staff in the Digital Engagement team and select committee offices to turn down future online engagement sessions due to the increased strain on resources and a lack of time to properly evaluate the many comments they may receive (Liaison Committee, 2015). These problems of managing resources and internal processes detract from the primary task of using digital methods to reach people who would not otherwise engage with the Parliament due to where they are located or time constraints. Therefore, it is important to find methods to break down the internal barriers to online engagement and provide a process which facilitates the use of online tools in parliamentary engagement.

The sections in this chapter each aim to answer one specific question which would usually be tackled through manually reading each comment using natural language processing. By automating this manual process, the value of data generated through online engagements can be maximised for parliamentary purposes. The chapter will proceed as follows; section 6.1 introduces the discussions to be analysed and provides preliminary textual analysis of the comments; section 6.2 will focus on using the meta-data of the discussions to uncover clues about the people engaging in these activities; while sections 6.3 and 6.4 delve deeper into the textual analysis to understand the topics which the participants find important and the

[26] Personal communication, Westminster

sentiments they express. The second half of this chapter (sections 6.5 and 6.6) focusses on the demonstration tests conducted with three select committees in the UK Parliament. I analyse how the use of a different platform, Discourse, which is built specifically for online discussions impacts the input citizens generate through their discussions and the way participants interact with each other compared to a social media platform, Twitter, which has been re-purposed to conduct online engagement activities.

6.1 A sample of parliamentary debates

I explore a sample of debates from various points in the year and on a range of subjects to gain a deeper understanding of who is communicating in these engagement activities, what subjects they find important, and how they feel about these subjects. The majority of these debates are held on a single digital platform (in this case Facebook, Twitter or a survey), however there is one debate which was discussed on multiple social media channels. This latter debate allows me to compare how the same subject is discussed on different channels. In doing so, I can explore how different channels can influence the way audiences react to the same issues.

All discussions took place throughout 2018 during the 2017-19 parliamentary session. To get a first sense of the debates, each of them was analysed based on most frequent words and word pairs (bigrams). This is visualised in a bigram network with the size of the points showing the word frequency and the thickness of the connections signifying the frequency of their occurrence. These discussions are summarised in Table 5, and show that the Fireworks discussion on Facebook and the Visit Visa survey had the highest number of comments at over 6000 each and also had the highest average comment lengths, while the Online Abuse discussion on Facebook had the fewest comments at 340, but had the third highest average comment length. This suggests there is not a direct correlation between the number of submissions and the substantialness of the submissions. Instead, the extent to which participants give their views could be more heavily impacted by the topic of conversation. The Digital Engagement team were involved in all discussions in this chapter, and the Petitions Committee and Web and Publications Unit (WPU) were also responsible for some of the discussions.

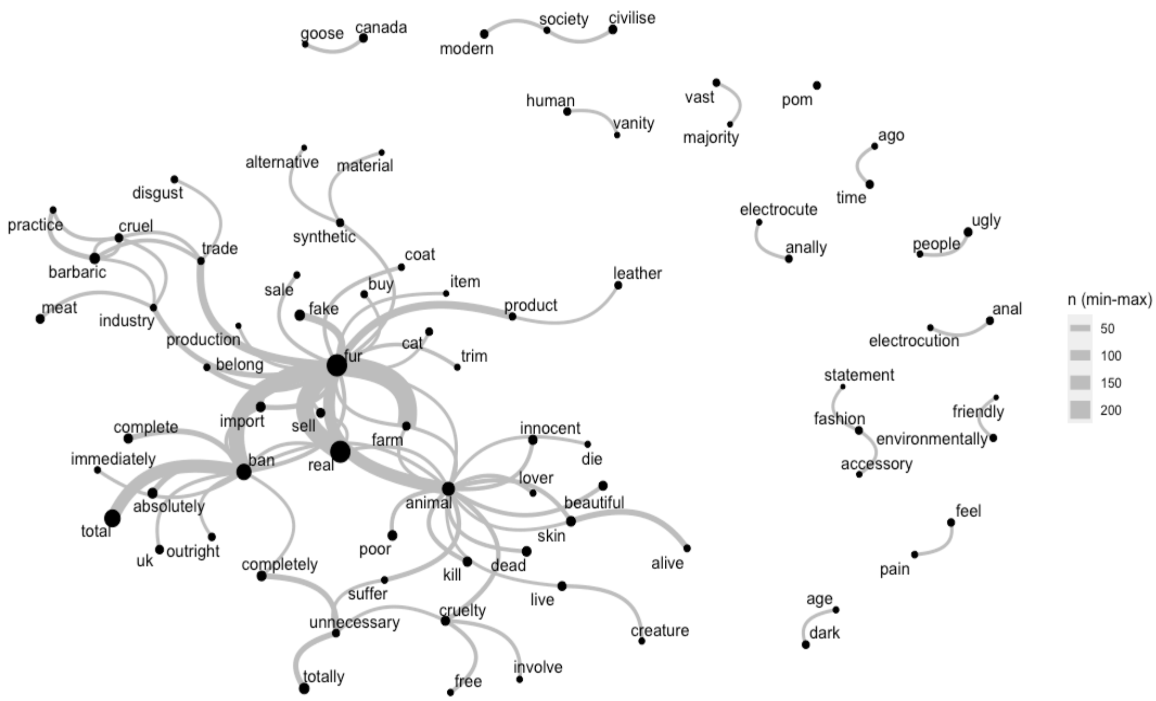
Table 5: Selection of digital debates

	Fireworks (FW) Debate		**Animal Fur (AF) Debate**	**Visit Visas (VV) Debate**	**British Sign Language (BSL) Debate**	**Online Abuse (OA) Inquiry**
Number of comments/ submissions	6012	2348	2794	6609	534	340
Mean comment length (words)	50	28	26	43	21	39
Engagement Channel	Facebook	Twitter	Facebook	Survey	Twitter	Facebook
Team Responsible	Digital Engagement		Digital Engagement	Petitions Committee/ WPU	Petitions Committee/ Digital Engagement	Petitions Committee/ Digital Engagement

Animal Fur

The animal fur debate recorded responses to the question "Should the sale of animal fur be banned in the UK?". This debate began as an e-Petition and was held as a digital discussion on Facebook on 18th May 2018, where a total of 2794 comments were posted (Ares, 2018). After reaching the necessary 100,000 signature threshold, the petition was debated in Westminster Hall on Monday 4th June 2018. The bigram network in Figure 31 displays the most common word pairs found in the discussion. The larger nodes signify words with a higher overall frequency in the discussion, and the thickness of the edges (connections) increases as the frequency of the bigram increases. This shows frequent words such as "ban", "fur", and "animal" are amongst the most frequent words each appearing over 1000 times. However, "cruel" and "barbaric" are also mentioned almost 300 times each. This already suggests that 1 in 10 of the comments posted used these negative adjectives in relation to animal fur. Looking further at the bigram network, word pairs "real fur", "total ban", "ban fur", "animal fur", and "faux fur" appear most frequently in this debate. These phrases are all related to the production of animal fur and are therefore to be expected. However, phrases such as "anal electrocution" which is a method used to kill animals whose fur will later be sold, shows a degree of subject-specific knowledge within some of the comments, different to the more emotionally-led "innocent animals" or "cruel practice". At first glance it appears this discussion will be relatively emotional and against the selling of animal fur with adjectives such as "suffer" and "pain", but also have details of the process of producing animal fur products.

Figure 31: Animal Fur Facebook Debate - Bigram network

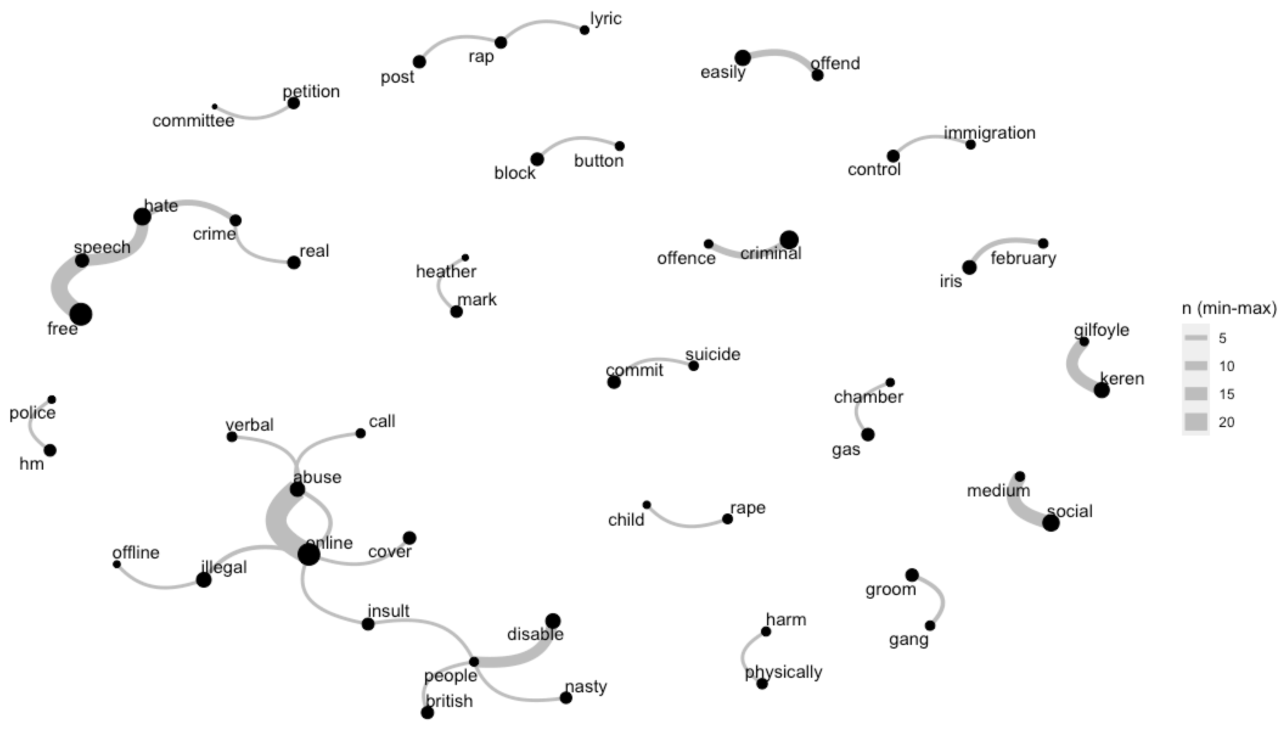


Online abuse

The discussion of online abuse of people with disabilities held on Facebook on 28th August 2018 was part of a larger inquiry by the Petitions Committee. This inquiry gained a large amount of press attention and was initiated through an e-petition started by British celebrity Katie Price who had a personal connection to the inquiry through her eldest son (Petitions Committee, 2018c). The petition was very successful gaining a total of 221, 914[27] signatures and was debated in Parliament on 29 April 2019.

This Facebook debate only featured 340 comments and is the smallest of all discussions explored in this chapter. The most frequent words (Figure 32) include “criminal”, “abuse”, “disable” and “online” and the bigram network also shows most participants were in favour of recognising online abuse of disabled people as criminal with word pairs such as “real crime” , “hate crime”, and “criminal offence”. However, the network also shows some participants had a less sympathetic view of online abuse with bigrams such as “easily offend”, “free speech” and possibly suggesting people use the “block button”. There was also a discussion surrounding the physical and mental effects of online abuse with “groom gang”, “physically harm”, “verbal abuse” and “commit suicide”. Therefore, even with a simple bigram network, I can already uncover several themes of discussion and develop insights into the specific feelings of the users who participated. This is a very useful feature for the Digital Engagement team who can use a bigram network analysis to have a preliminary understanding of some of the themes and feelings of participants in a discussion without yet reading any individual comments.

Figure 32: Online Abuse Inquiry - Bigram network

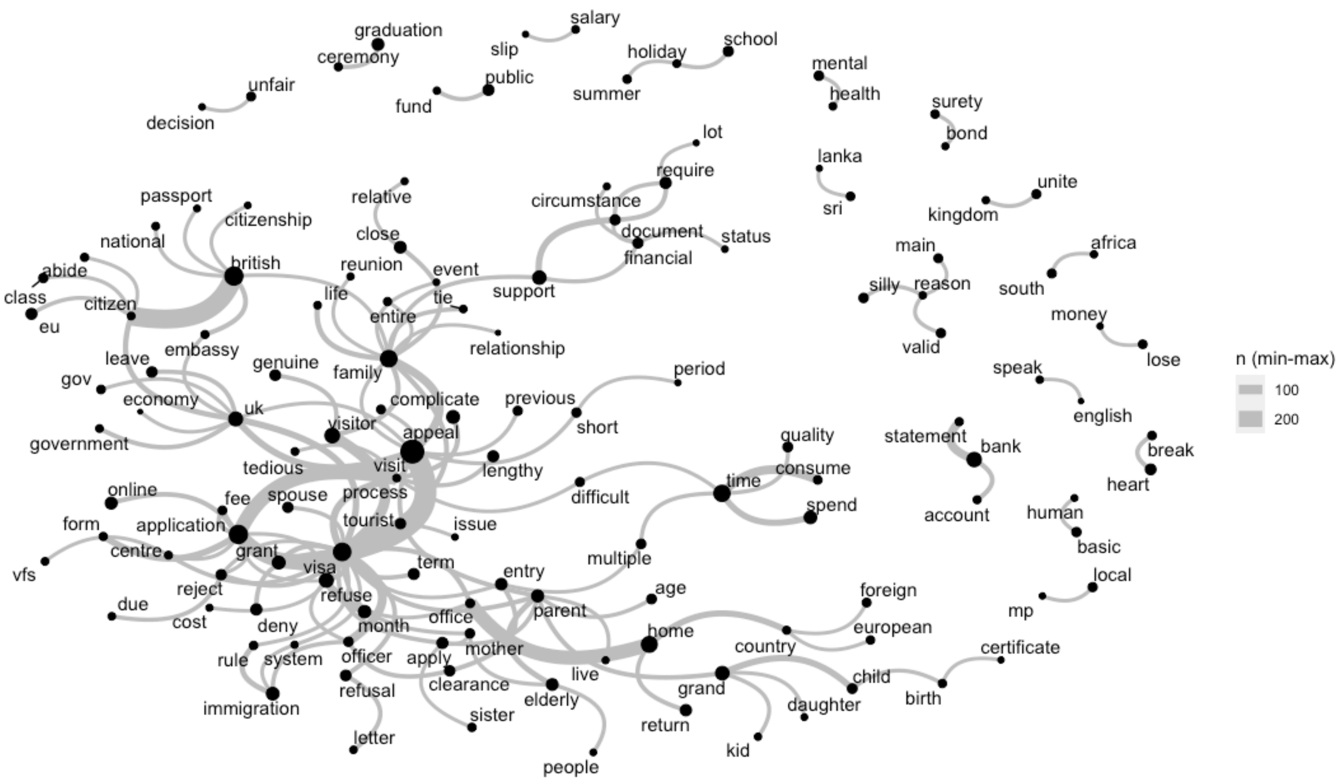


[27] https://petition.parliament.uk/archived/petitions/190627

Visitor Visa

The Visitor Visa inquiry was held by the Petitions Committee in 2018 and aimed to gather evidence from members of the public who had been affected by the application for a visit visa. The textual data comes from a survey sent to members of the public in June 2018 , however the e-petition “Introduce automatic approval of visit visas for families of British Citizens” was also debated in Westminster Hall on 9 July 2018 (Petitions Committee, 2018b). Although this petition received 71,178 signatures which is under the 100,000 threshold for a debate, it was debated in Parliament due to the Petitions Committee inquiry. The most frequent words in the bigram network (Figure 33) reveal that “application”, “process”, and “refuse” are common words used by submissions to the inquiry. The bigram network also shows that “home office”, and “time consuming”, join “visit visa application process” are the most frequent word pairs. This suggests that there is a general frustration with the process of applying for the visa, both with the issuing authority (home office) and the length of time as indicated by “6 months”. Certain countries are highlighted such as “Sri Lanka” and “South Africa”, and some family occasions that may have been missed as a result of the visa rejection such as “family reunion” and “graduation ceremony”.

Figure 33: Visa Inquiry - Bigram network



Fireworks Debate

This discussion was held on Facebook asking the question “Do you think the use of fireworks needs to be more regulated?”. This was in response to an e-petition and was debated in Parliament on Monday 29th January 2018 (Petitions Committee, 2018a) . During the Westminster Hall debate of the e-petition, a Twitter hashtag was created by the Petitions Committee and used to follow proceedings. Over the course of 48 hours, this hashtag generated 2348 tweets which used the #Fireworks hashtag. This allows me to compare the two debates on different platforms but on the same issue. Additionally, Twitter data allows for a social network analysis of participants in the Twitter debate, which gives additional insights. A detailed analysis of the network of retweets to explore how this discussion was comprised of separate groups of users discussing different issues, users will be explained in section 6.2.1.

Figure 34 shows the bigram network including frequent words used in the Twitter discussion and Figure 35 contains the same information for the Facebook discussion. The figures show that the Facebook discussion was much more specific to the intended discussion topic than the Twitter discussion which had word pairs such as “australia day”, and “enjoy india” among the most frequent. This Twitter discussion also contained the phrase “curso diseo basico para redes sociales” which translates to “basic design course for social networks” showing that other discussions using the #Fireworks hashtag also not relevant to the e-petition were also ongoing during this time. This is because the hashtag itself is very general. The Petitions Committee has shown in the past a somewhat inconsistent approach to using hashtags, sometimes using existing general ones, and sometimes generating hashtags that are very bespoke. The Facebook bigram network (Figure 35) shows that the most frequent pairs of words used is “organised/professional/regulate displays” which suggests that participants were advocating organised firework displays rather than spontaneous amateur displays. This network also shows concern for animals with “poor dog”, “farm animal”, “domestic animal”, and “animal suffer” each appearing over 50 times in the comments. The bigram network allows for a preliminary overview into the topics raised by participants in the discussion. This will be explored further using topic modelling algorithms in section 6.3.

Figure 34: Fireworks Twitter Debate – Bigram network

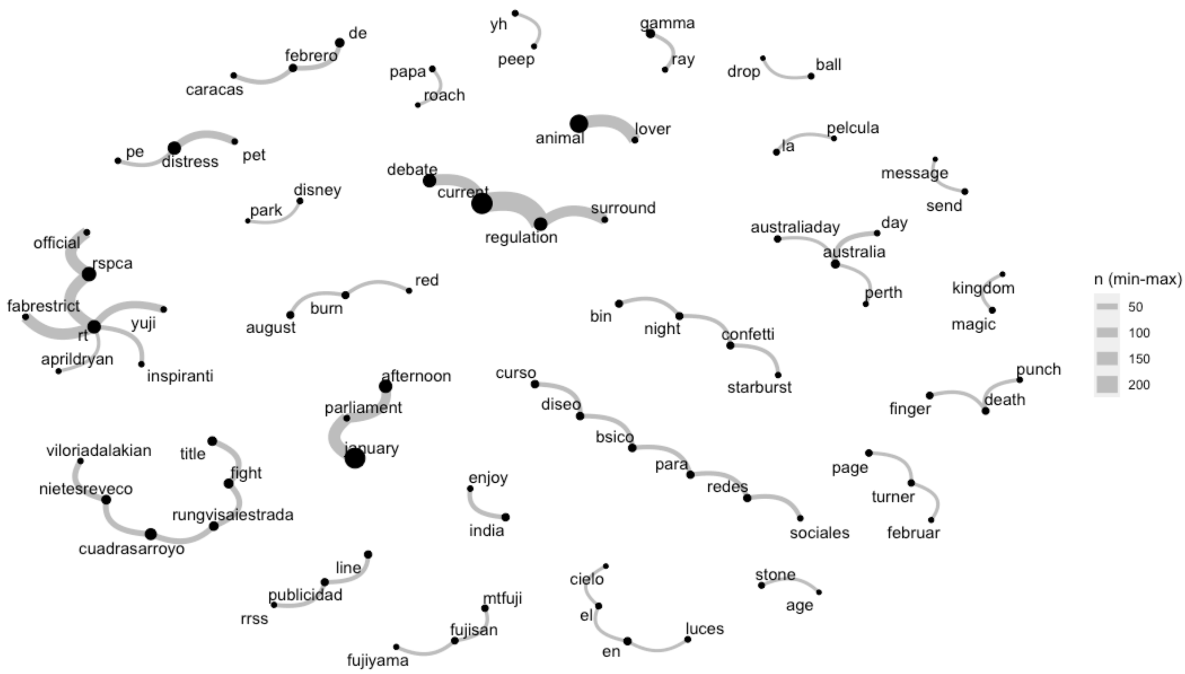
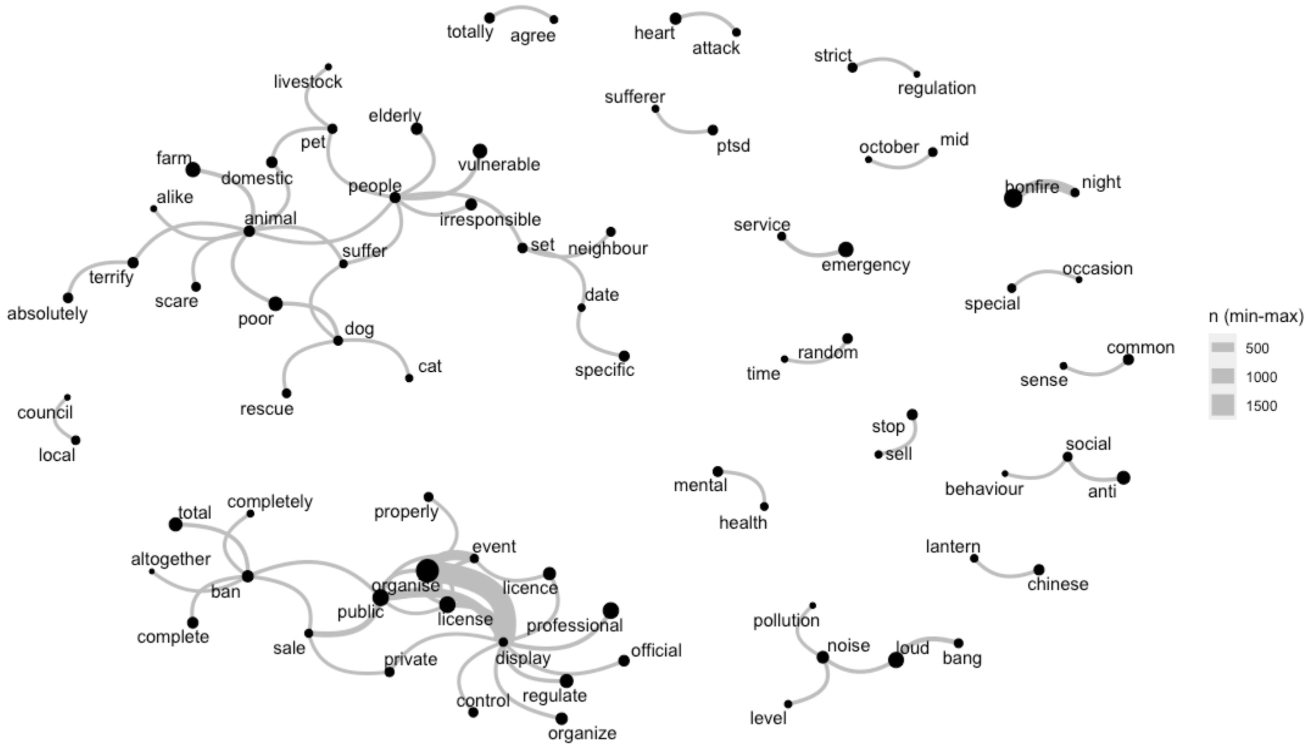


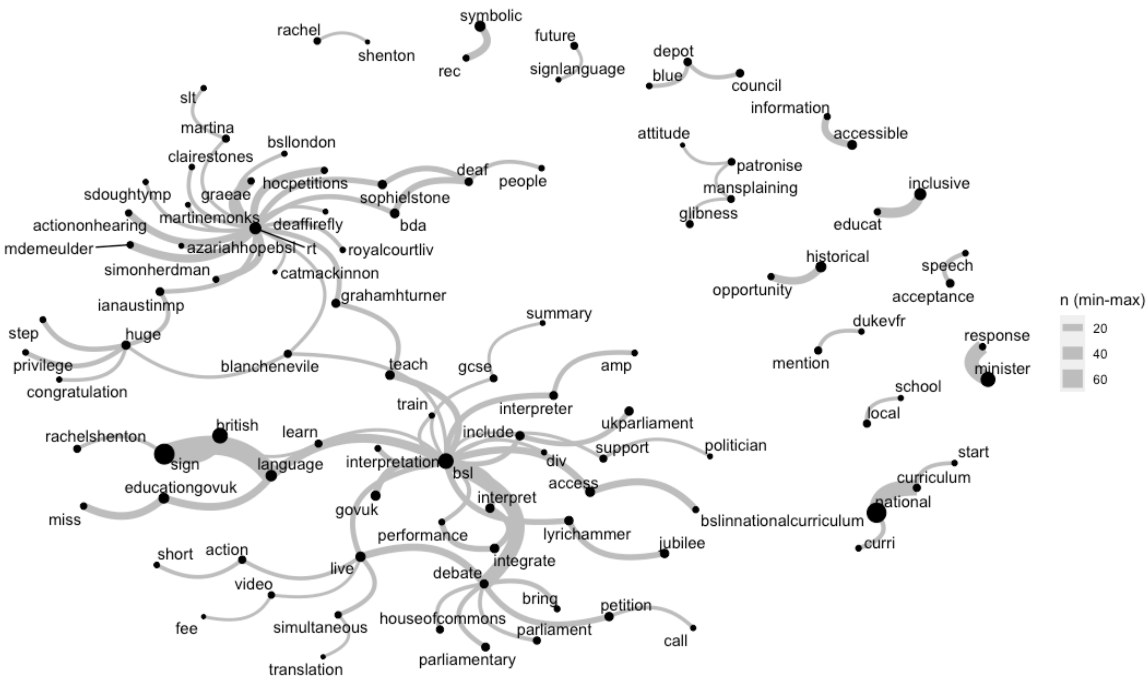
Figure 35: Fireworks Facebook Debate - Bigram network



British Sign Language

The British Sign Language (BSL) Twitter debate followed the e-petition “Make British Sign Language part of the National Curriculum”. The tweets were recorded using the #BSLDebate hashtag during the Westminster Hall debate on 5th March 2018 (House of Commons Library, 2018). The debate was held during British Sign Language Week and featured a live BSL interpreter in Westminster Hall for the first time. The top 10 words in Figure 36 show expected words such as “bsl”, “sign”, and “deaf”, but also make reference to the “live” interpreting of the Westminster Hall debate. However, “patronise attitude” and “mansplaining” also appears in the most frequent word pairs on Twitter suggesting not all users were happy with the discussion. The bigram network also reveals that several Twitter accounts received a lot of retweets, for example “rt dawnbutlerbrent” and “rt graeae”. There was also a focus on ‘inclusive education’ and ‘historical opportunity’ related to the live interpreting of the debate. This could suggest users were welcoming to the efforts Parliament was making to be inclusive to more members of society wanting to participate in the discussion.

Figure 36: BSL Twitter Debate - Bigram network



6.2 Who are the participating citizens?

One of the primary questions when conducting engagement sessions online is getting an understanding of who is participating in the activities. One aim of parliamentary engagement is to reach a varied audience (Liaison Committee, 2015) and in certain cases it is possible to ask for background information during the activity, for example in a survey. However, another aim of parliamentary engagement is to go where the people are and use existing communities. In an offline setting this can be achieved through outreach exercises which take staff across the country to people who live far from London. Some committees have also held focus groups or travelled to community groups who have lived experience of a particular issue. During these offline engagement sessions, it is easy to see the demographics of participants regarding age, race etc, how they interact with each other, and get information on who they are. However, in an online setting, going where the people are and using existing communities means social media and online networks. Where this has the benefit of allowing many people to interact irrespective of their physical location, it has the disadvantage of being more difficult to know who is engaging. For example, on Facebook, users can decide how much personal information they are willing to provide, and it is increasingly difficult to get this information when exporting discussion data.

Therefore, to answer the question about who is participating I use a mixture of social network analysis of Twitter data, activity data from Facebook discussions, and socio-demographic inference from textual comments to build a picture of the people behind the comments.

6.2.1 Who are the key players? Network Analysis of Twitter interactions

A disadvantage of using certain types of social media is the limits on data exports. Section 5.3.3 included analysis of the network of interactions between Facebook users and the House of Commons Facebook page. I saw that users interacted more with certain topics over others, and could be grouped based on their interactions with similar posts. However, these insights were only possible due to the ability to export network data from Facebook using the Netvizz application (Rieder, 2013). Unfortunately, this is no longer possible on Facebook, however it is still possible to analyse the network of interactions for discussions held on Twitter. In my sample of debates, this can be done for Fireworks and BSL discussions. These Twitter discussions were each captured using a specific hashtag decided by the Petitions Committee. The data was then collected over 48 hours – starting 24 hours before the Westminster Hall debate which takes place on Mondays 16:30, and finishing 24 hours after the debate. Centrality measures and community discovery algorithms are used to evaluate interactions between users. The nodes represent Twitter users and the edges a retweet interaction. A retweet can include both a mention of the original tweet and the user's response to that tweet. The retweets are also denoted by a "RT" icon which can be easily extracted from the raw data. The nodes represent Twitter users and the edges a retweet interaction.

Different measures of node centralities can shed light on the node's position and influence over a network. Figure 37 shows the network of Fireworks Twitter debate with the size of the nodes ranked by their degree. A high degree indicates that a node has many connections with other nodes, and vice versa. In the Fireworks network I can see that the @RSPCA_official – a UK animal charity, @Yuji_48 – a Japanese photographer, and @MWOBS – the Mount Washington Observatory have the largest nodes and therefore the highest degrees. I can also see that these nodes are at the centre of clusters of other users, almost completely distinct from other clusters. The small nodes (with low degree values) have all retweeted a tweet by the central node of their cluster and suggest individual clusters of users who have little to no interaction with others. The social network based on the Fireworks debate on Twitter in Figure 37 is an example of a *broadcast network* as identified in Figure 11, where users are retweeting a single account but without much interaction between themselves. This suggests that although these users are actively participating and viewing the information tweeted by the main accounts, there is not much discussion and these users could just be passive participators who are scrolling their Twitter feed rather than actively engaging.

Furthermore, the hashtag used in this case was #Fireworks and attracted a lot of users who had no relation to the Westminster Hall debate or e-petition at all. This could account for the number of isolated clusters as they were distinct communities with no relation to each other, but just happened to use the same hashtag at the same time in their tweets. For example, @Yuji_48 tweeted a photograph of a fireworks display in Japan which received 73 retweets[28], while @MWObs tweeted a photograph of fireworks seen at the base of a mountain which received 64 retweets[29]. This raises a difficulty in using Twitter as a data source for public engagement if the internal process is not clarified beforehand. Using a very generic hashtag such as #Fireworks leaves the discussion vulnerable to derailment by users who have nothing to do with the intended discussion, as was the case for Fireworks. A more specific hashtag such as that used in the BSL discussion reduces the likelihood for other users to mistakenly hijack the discussion and cause confusion for the parliamentary team leading the engagement activity. In most analysis, anomalies or irrelevant data would be omitted from results, however, to do that at this stage would create a false perception of the discussion. The aim of this chapter is to keep as much of the original debate environment the same, use the exact same data as the parliamentary teams would have, but to use natural language processing and social network analysis to gain insights into the data. The deliberate inclusion of this data highlights where digital engagement activities can fall short on reaching the intended audience and make a

[28] https://twitter.com/Yuji_48/status/957960373461647361
[29] https://twitter.com/MWObs/status/957457226272264193

discussion appear more disjointed and the users more unconnected than they really are (for example in Figure 37).

Figure 37: Fireworks Twitter Debate – nodes weighted by degree centrality

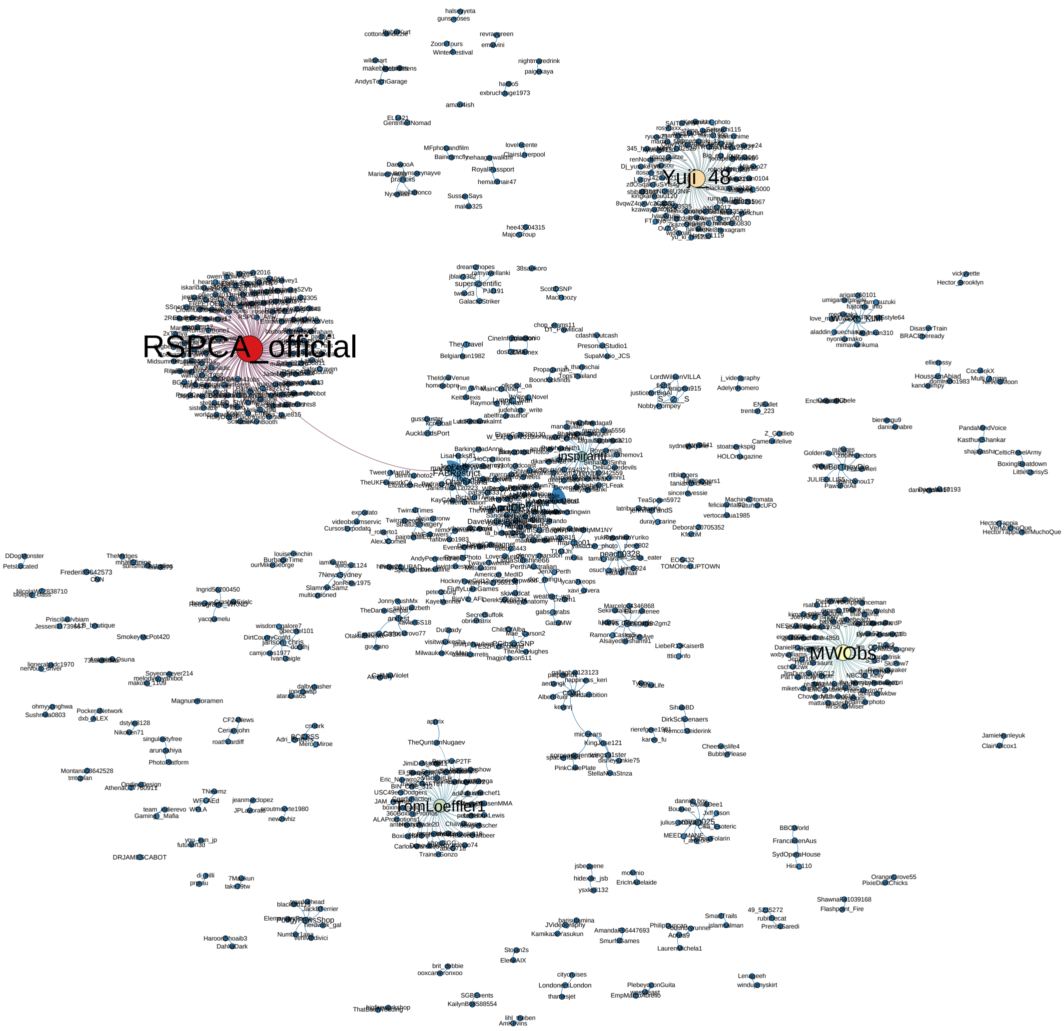
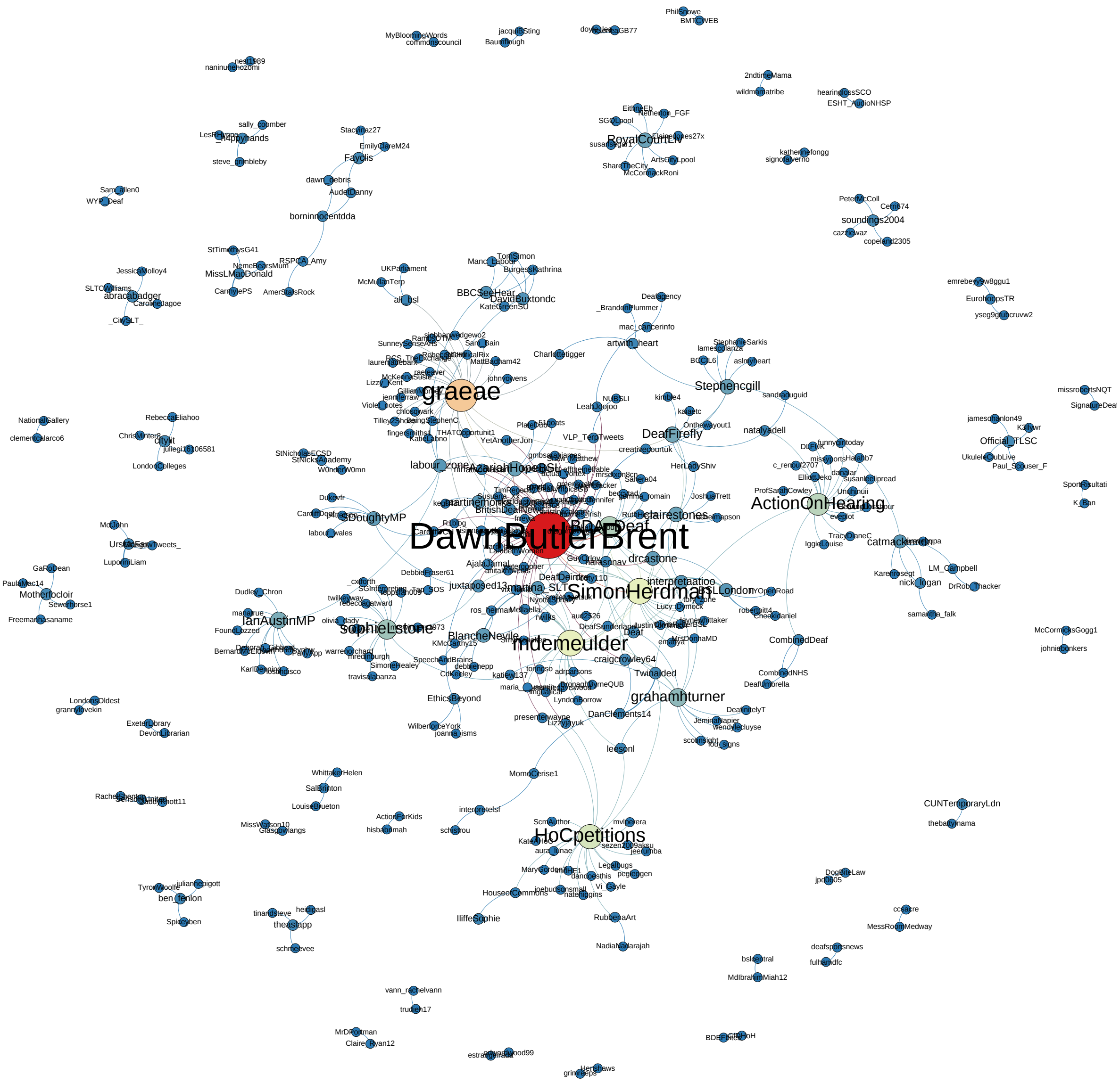


Figure 38 shows the same degree ranking for BSL Twitter discussion. At first glance, it has already a very different structure to Fireworks, with fewer isolated large clusters. This suggests the pattern of retweets among users using this hashtag is less centred on individual communities and has more interconnectivity. Users @DawnButlerBrent, @graeae, @SimonHerdman, @mdemuelder, and @HoCpetitions had the highest degrees and the largest nodes. Although this network is generally more connected than the Fireworks one, there still exist clusters of users centred around a single node with only a few connections to nodes outside of the cluster. There also exist pairs of users completely outside the core centre of the graph. These more isolated communities are mainly comprised of between 2 and 5 users who have either retweeted each other (in the case of 2 users) or several users have retweeted a single account (in the case of 3+ users). As mentioned before, this Twitter discussion used a much more specific hashtag of #BSLDebate which appears to have led to a more connected social network with participants interacting much more thoroughly with each other than with the Fireworks Twitter discussion.

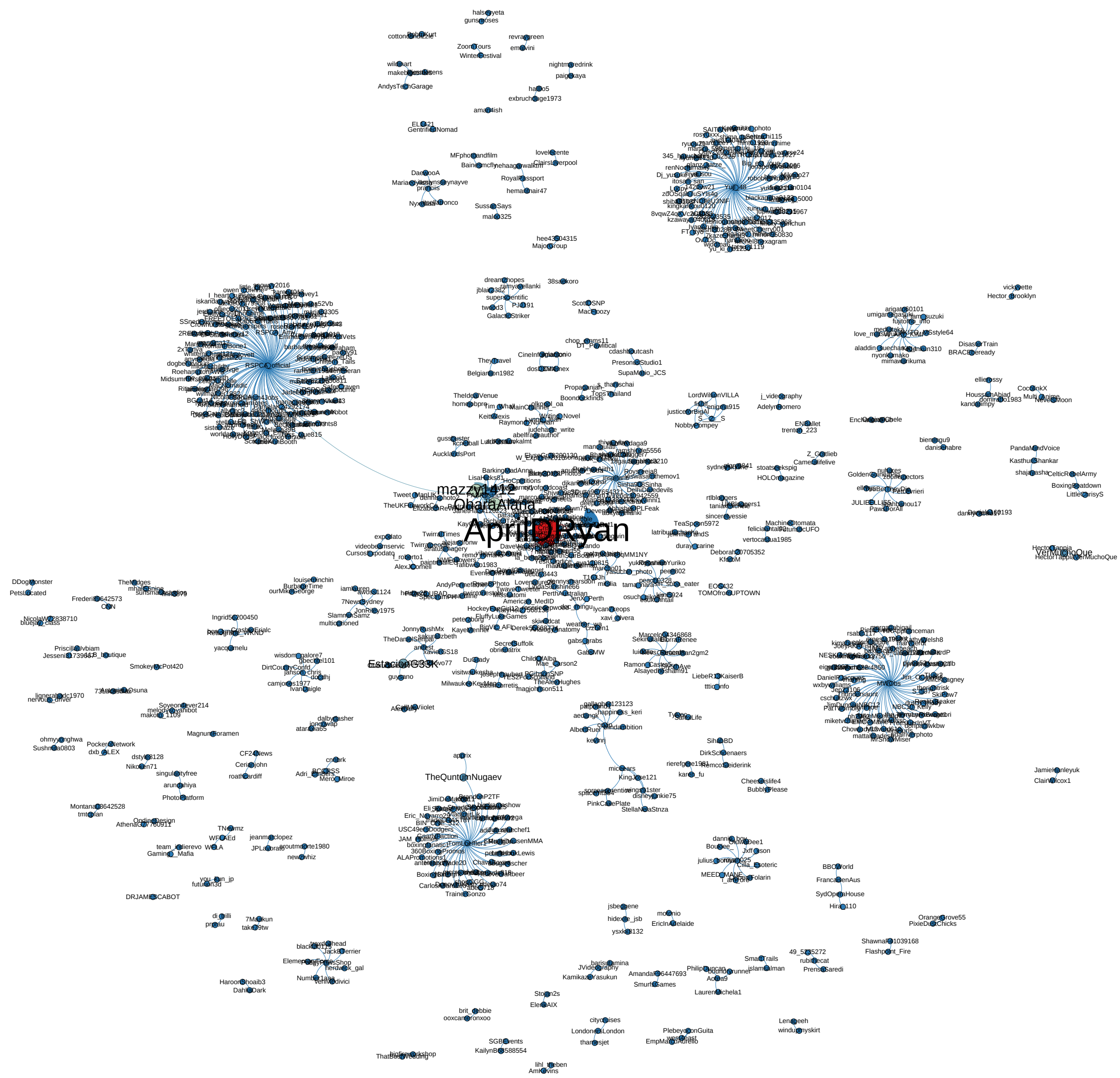
Figure 38: British Sign Language Twitter debate – nodes weighted by degree centrality



Where degree centrality takes into account the raw number of connections of a node, betweenness centrality takes into account the number of edges that are crossed by that node. This can provide clues on the most important node rather than the most popular node. Plotting the same Fireworks network as Figure 37, nodes in Figure 39 are large if they have a high betweenness centrality. Figure 39 shows that the users who had the largest degrees have small betweenness centrality values, and it is user @AprilDRyan who intersects the most connections by far, followed by @mazzy1412, and @ObaraAlana. This means that although @AprilDRyan was not retweeted by many users, she acts as a bridge or connection node between many other nodes, which suggests she has a varied Twitter following within this particular hashtag.

Looking closer at this user reveals she is a White House Correspondent and CNN Political Analyst based in Washington, DC[30]. Using Twitter's search function filters[31], she made no specific reference to fireworks or the use of the hashtag, however her followers appeared to have used this hashtag in relation to her. This again could be a consequence of the very general #Fireworks hashtag including users and interactions with very little to do with the intended e-petition and Westminster Hall debate.

Figure 39: Fireworks Twitter Debate – nodes weighted by betweenness centrality

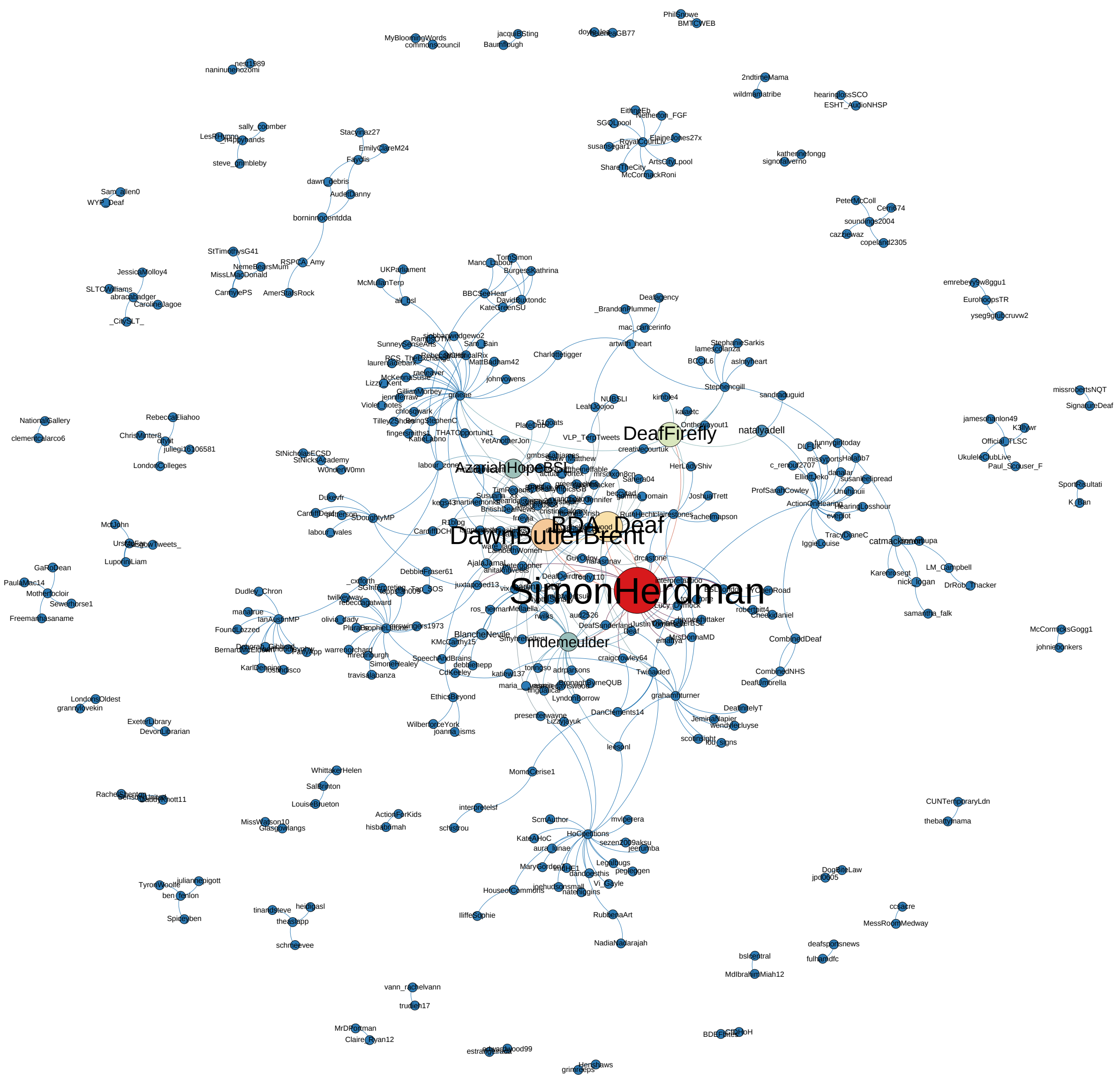


Within BSL, the betweenness centrality rankings in Figure 40 are different to the degree rankings (Figure 38) but to a lesser extent as Fireworks. This time @SimonHerdman who had

[30] https://twitter.com/AprilDRyan
[31] from:AprilDRyan since:2018-01-26 until:2018-01-31

a moderate degree, has the highest betweenness centrality suggesting he holds a similar position in the Fireworks network to @AprilDRyan in the BSL network in crossing many different connections within this discussion. He is an activist and filmmaker for the deaf community[32] and could therefore have a higher 'importance' and influence within the network. This suggests the BSL discussion had participants who had a lived experience of the issue and could make meaningful contributions to the discussion. This is a characteristic the Digital Engagement team would favour in their online engagement activities as it is indicative of a meaningful discussion which would lead to valuable contributions that they can pass on to MPs.

Figure 40: British Sign Language Twitter debate – nodes weighted by betweenness centrality



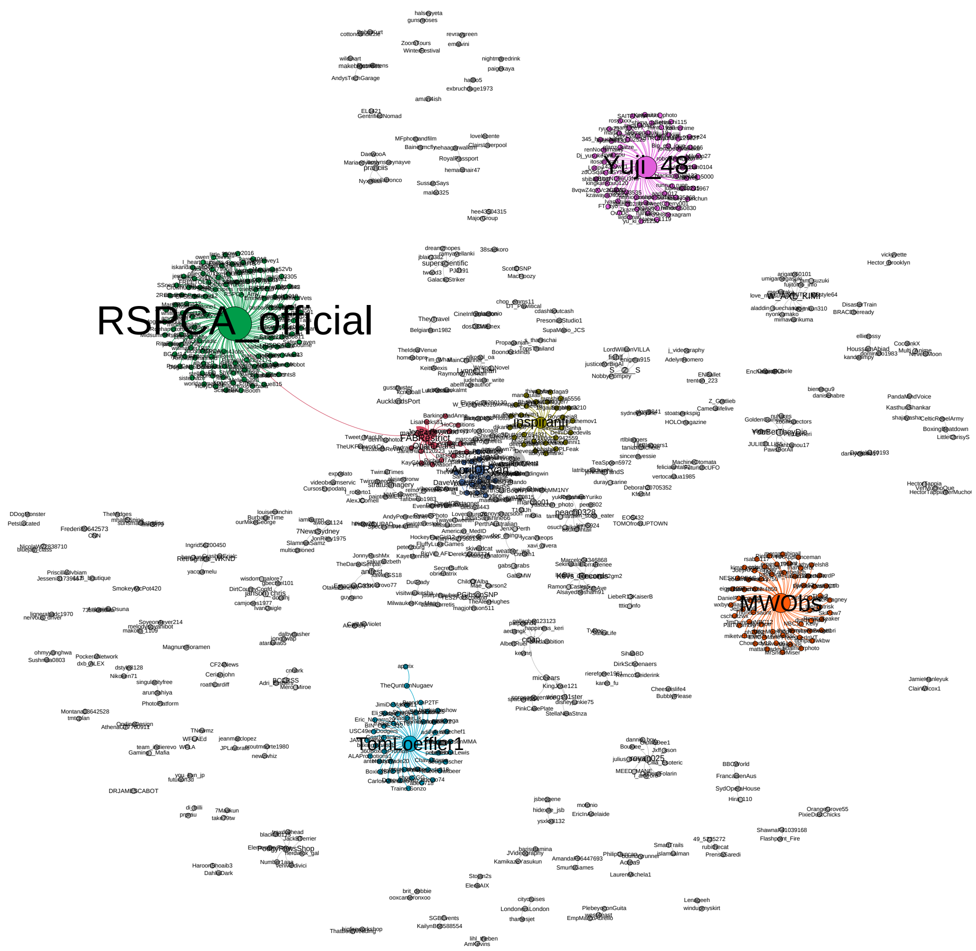
Measuring centralities in this way allows me to have a picture of the important nodes in the network and how they interact with each other. Figure 37 and Figure 39 show that the

[32] https://twitter.com/SimonHerdman

Fireworks Twitter debate was a very disjointed discussion with users often interacting exclusively with users within their own cluster, however this could be due to the vagueness of the hashtag. I see that the primary actors in the discussion, both based on popularity and importance, did not include the @HoCpetitions account who initiated the discussion. This will be explored further in the next section. In contrast, Figure 38 and Figure 40 show a much more connected Twitter discussion for BSL, not as heavily centred on individual users.

The centrality measures provide some insight into which nodes have the most influence in a discussion but do not show the graph-level statistics of the network. These are important to understand how all the nodes in the network interact and uncover the presence of any community clusters. The Louvain modularity algorithm is used to assign clusters to the nodes (Kolaczyk, 2009), and a higher value for modularity (between 0 and 1) indicates the nodes in the network can be well grouped into distinct clusters based on their edges. In Fireworks, the modularity was 0.93, and each of the 7 nodes with the highest degree are the focal point of their own cluster. It can be seen that although @HoCpetitions was leading the debate, they are actually in a different cluster (red) to the @RSPCA_official account (green) which was the most retweeted (Figure 41). The @HoCpetitions account was only retweeted by one person (@JenPadleywood) who retweeted a few other people including the @Mazzy1412 account who then retweeted @RSPCA_official. Therefore, @HoCpetitions had 3 degrees of separation from the main hub (15% of users).

Figure 41: Fireworks Twitter Debate – nodes weighed by degree centrality and clustered Louvain method

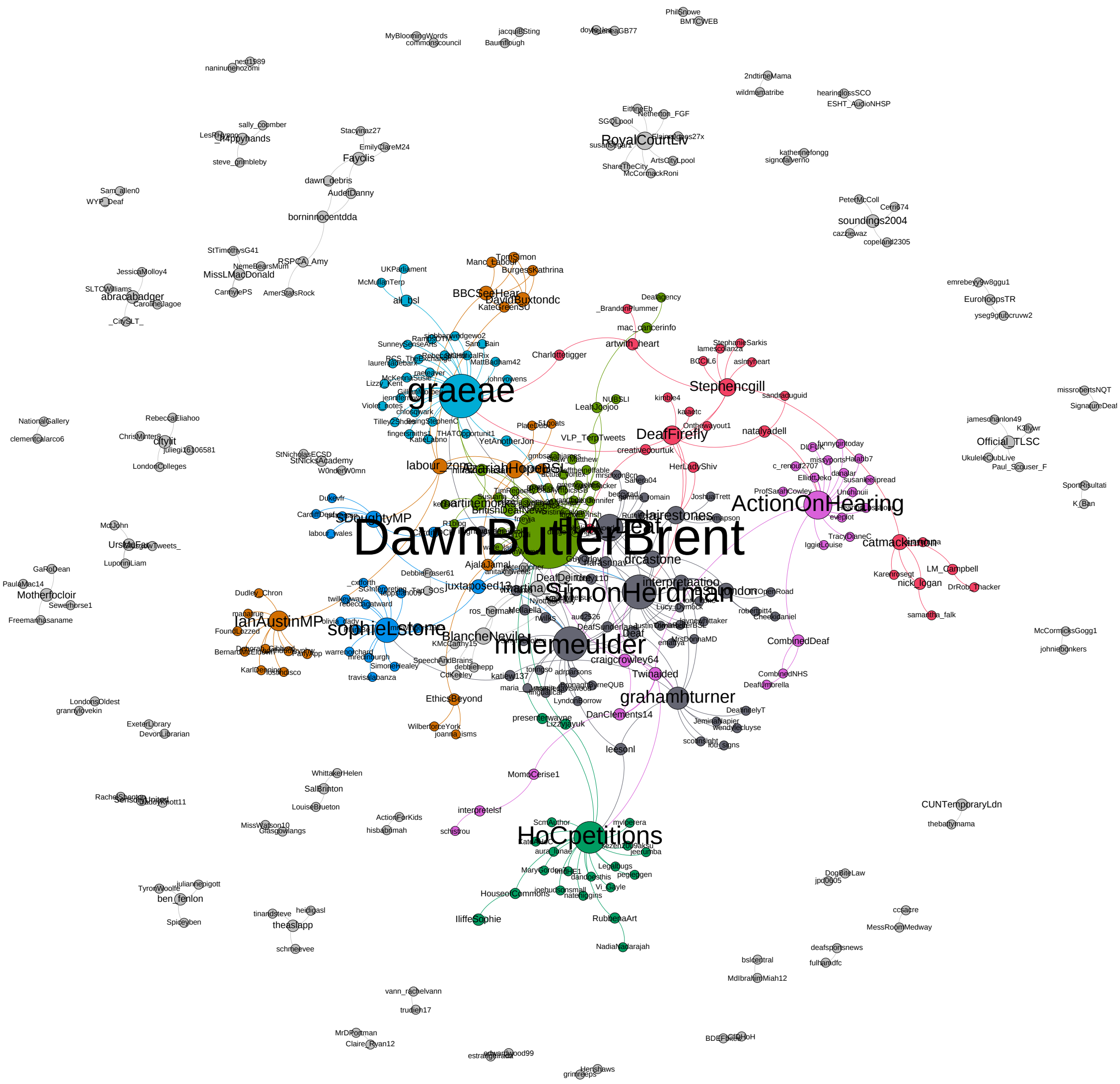


The BSL debate clusters (Figure 42) are not as disconnected as the Fireworks. There is evidence of nodes from different clusters interacting with each other as well as with the nodes in their own cluster. This means users were not as divided in terms of only seeing or retweeting tweets that were made in their own circle. @HoCPetitons also has a much stronger role in this discussion with a relatively high degree, and direct connections to other clusters. I can also identify some individual users who act as bridges between clusters, or cutpoints. User @YetAnotherJohn is the sole connection between the light green @DawnButlerBrent cluster and the light blue @graeae cluster, and @Twidaided is the sole connection between the dark green @HoCPetitions and the purple @ActionOnHearing. This suggests these two users in particular retweeted tweets from completely separate communities and could indicate they have a varied list of users they follow, allowing them to see and interact with tweets from a range of accounts.

In addition to @DawnButlerBrent, there were also two Labour Party MPs involved in the discussion, @IanAustinMP and @SDoughtyMP. Both were centres of their own clusters of users however the former had a larger number of retweets from more users. Furthermore, neither of the MPs interacted with the @HoCPetitions account, despite being a part of the same institution. This could be because they were unaware of the Petitions Committee's involvement with the debate, or because their main concern was to interact with their own followers who are most probably their constituents. This shows another difference between party political engagement by political figures and non-partisan engagement by select committees which was also raised in Chapter 5. Politicians use Twitter to promote their own ideas and can freely express their political views with their followers, however Parliament always aims to represent the views of many within the political process.

Both Fireworks and BSL Twitter networks called on different charities related to their issues (RSPCA and Action on Hearing respectively), however Fireworks was much more dominated by the charity's involvement than BSL. The difference in network structure may have many different reasons. The two discussions ranged heavily in terms of the number of users using the hashtag with 2348 for Fireworks and 534 for BSL. Furthermore, the choice of hashtags may have made #BSLDebate more specific and therefore attracting a more informed set of users than #Fireworks. The BSL Twitter discussion was also part of an inquiry by the Petitions Committee and its Westminster Hall debate had simultaneous sign language interpreting for the first time in Parliament's history (Parliamentlive.tv, 2018). The Fireworks discussion is a topic of many submitted e-petitions and as a result may not have had the same levels of awareness. The higher levels of involvement of the @HoCpetitions account in BSL may have also had an influence on the types of users interacting.

Figure 42: British Sigh Language Twitter debate - nodes weighed by degree centrality and clustered Louvain method



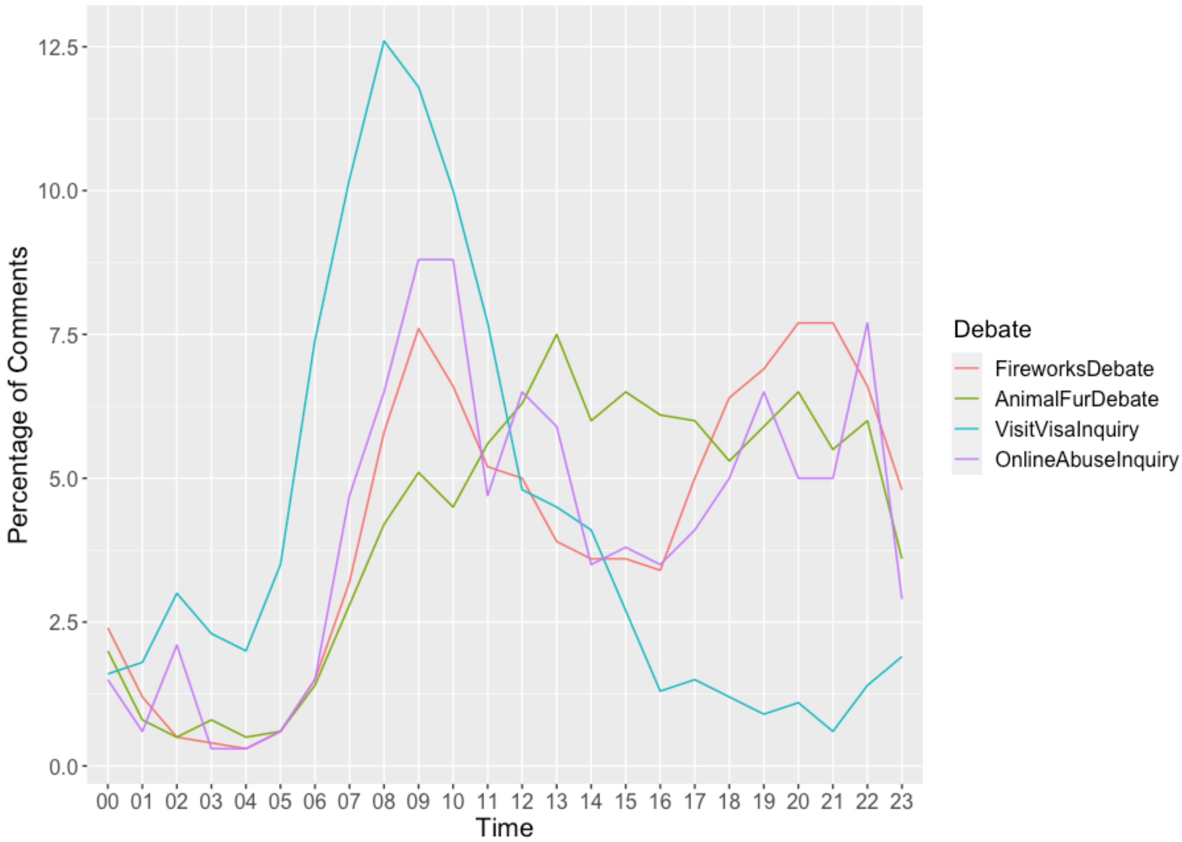
6.2.2 When are they most active?

Taking a discussion online has many advantages including allowing the public to participate from where ever they are, and whatever time is most convenient for them. In a parliamentary setting, this helps the institution become more open in their engagement and reduces the external barriers to engagement for citizens such as travelling to Westminster or taking time off work. While this project primarily concentrates on the analysis of text and network structures of digital discussions, valuable information also lies in the temporal meta-data of these online engagement activities. If working under the assumption that people will generally participate when they have some free time, through analysis of the timings of discussions, I can gain insights into the daily patterns of the participants in my sample.

The most popular times of day to post a comment differ strongly between the different debates. In Fireworks and Online Abuse digital discussions, one of the most popular times of the day for users to post was during the morning followed by a decline in mid-late afternoon, and these topics experienced another peak of activity towards the end of the day (Figure 43).

This pattern coincides with the post-rush hour routine with peaks just after the morning rush hour (07:00-09:00) and the evening rush hour (16:00-18:00). This suggests the people commenting on these debates follow a general 9-5 working routine, leaving them available to participate in the discussion in the evenings after work and in the morning, perhaps during their

Figure 43: Percentage of comments posted to digital debates per hour



commutes.

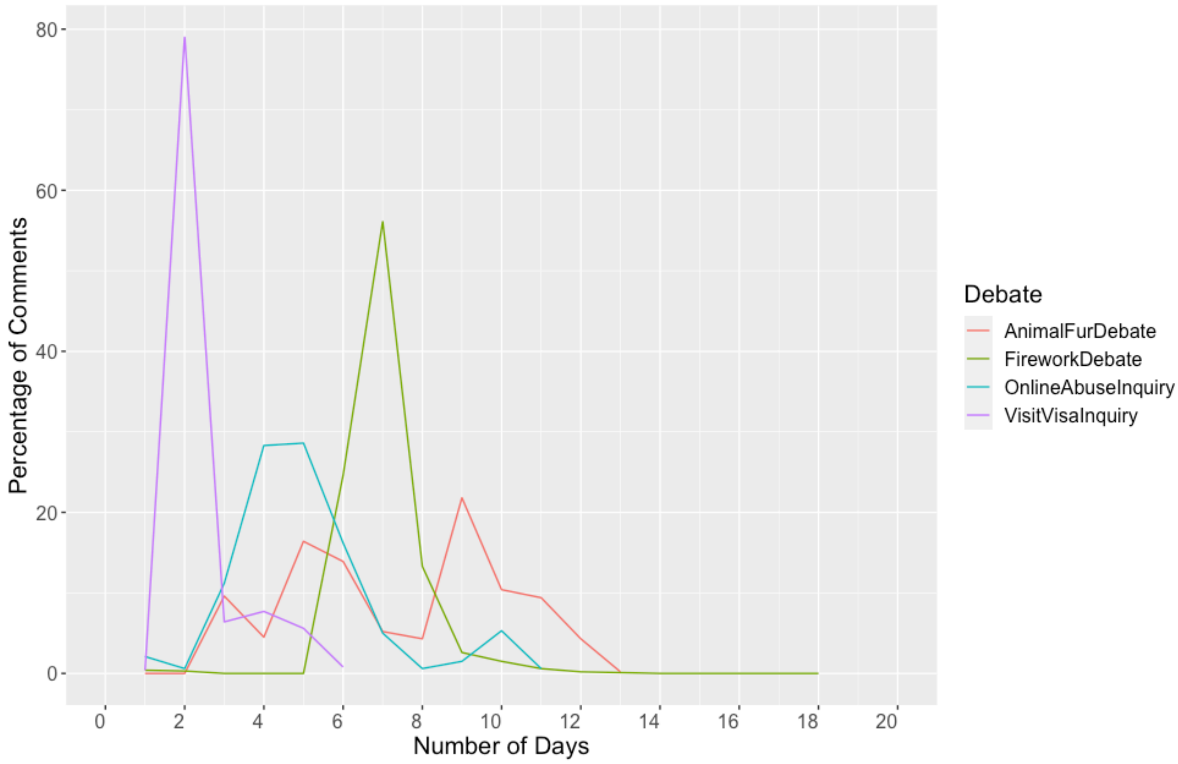
Similarly, the Visit Visa debate saw a large peak in activity in the morning between the hours of 06:00 and 11:00 but then witnessed no real activity for the rest of the day. This could be due to the channel the comments were recorded through being a survey rather than social media. Relating back to the different dimensions of engagement, the Visit Visa survey more resembled a consultation activity where participants were unable to see any submissions from each other but simply answered a set of three questions on the topic of visitor visas. This differs from a discussion activity where participants were able to read and respond to other participants' comments on Facebook. In practice, this also means that once users completed the survey there was no reason for them to revisit it as their participation was complete. This can be witnessed in the single peak of activity within the Visit Visa survey compared to the various peaks and troughs of activity seen in the Animal Fur, Fireworks, and Online Abuse discussions. This supports existing research which finds that a user's behaviour is influenced by being able to see the participation of others (Hale *et al.*, 2018).

Unlike Fireworks and Visit Visa, in the case of the Animal Fur debate, the public posted a during much wider range of times. The only time of day where a considerable drop in commenting was found is between midnight and 06:00. This implies that the people participating in this debate are less restricted to the general 9-5 routine as demonstrated in Fireworks and Online Abuse discussions. Instead, anytime of day appears to attract around the same number of comments, apart from a slight peak around 13:00.

Moving now to a dates perspective, the vast majority of the Fireworks comments were posted between the 27th and 29th January or between the 6th-8th day of the discussion. Visit Visa, Online Abuse, and Fireworks have a similar pattern for their date trajectories in that there is a clear peak of activity around a few consecutive days and the trajectories are clearly right-skewed, with greater activities at the start of the digital discussion (Figure 44). This can be expected for Visit Visa which was distributed as a survey to a set number of participants and with a limited time frame to complete. However, Fireworks and Online Abuse discussions were held as a Facebook digital discussion, starting with a House of Commons digital discussion card, which was available for a longer period of time and in a more informal format. This could suggest the Fireworks debate was very active and emotive causing many people to comment and respond very quickly. However, this may also suggest that they did not take time to think about their comments, or similar types of people were commenting – perhaps those who first saw the post. Topic and sentiment analysis in sections 6.3 and 6.4 will explore this further. The Fireworks debate was also shared by an animal rights charity which most likely caused the increase in comments and the small range in dates due to a large number of people viewing the discussion in a short period of time (once it had been shared by the charity).

The majority of the Animal Fur comments were posted on the 28th May 2018 - 9 days after the discussion card was posted by House of Commons on their Facebook account profile page. Whereas in the Fireworks debate, Online Abuse debate, and Visit Visa inquiry, most comments were made on one day there appears to be a greater variation of times when the Animal Fur comments were posted with peaks and troughs suggesting this was not a consistent discussion. The 22nd, 24th, and 28th of May (3rd, 5th, and 9th days) experienced between 250 and 600 comments (10%-20% of total comments) each day, while other days in between received barely 100 comments (less than 5%). This pattern could be a result of the subject matter of the discussion, the different types of promotion used in the Facebook post, people's schedules, or could be indicative of foul play or bots in the discussion. For example, users that post many times in a short space of time could be indicative of a bot, and a discussion being heavily promoted could result in a large number of different users viewing and interacting with the discussion. Overall, by examining the times and dates where comments were posted, I can gain some understanding of the types of users participating in these discussions or at least learn something from their daily patterns such as whether they work 9-5 hours, or more unsociable hours. I can also explore how the different discussions can attract different users in terms of how they chose to respond. Some discussions appear to have had participants returning day after day to continue the conversation, whereas others were much more based on users seeing the post, making a contribution and not returning to it again. These insights can be used for the Digital Engagement team to evaluate how certain discussions can be influenced by the way people post in them, for example the sentiments or different topics raised in each discussion. The next section will look a little closer at the socio-demographic makeup of the participants in these discussions.

Figure 44: Percentage of comments posted to digital debates per day



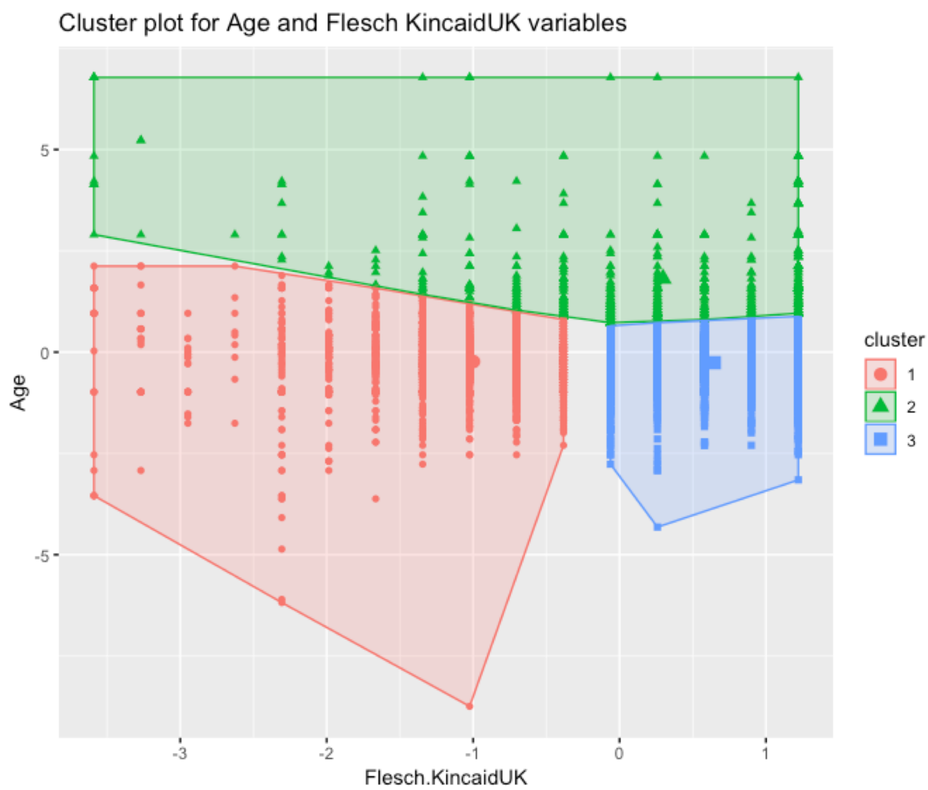
6.2.3 What is their socio-demographic background?

We have so far explored how different users can influence a discussion on Twitter which gives me a little insight into who they are and how they behave online. Whereas section 6.2.1 focussed on the insights gained through social network analysis and section 6.2.2 explored how the different times of posts can provide clues on who the participants are, this section looks at what other information I can obtain through more generic data analysis and natural language processing methods.

In my final method of understanding who is participating I attempt to infer socio-demographics from social-media data. Section 3.2.4 introduced a methodology for inferring socio demographic information from the linguistic clues users leave behind in their comments. This involves analysis of the prevalence of certain Part-of-Speech (POS) tags and combines this with the approximate UK grade level which is required to understand a piece of text, in my context a given user comment. To estimate the demographics of the users participating in my sample of discussions, the indicators for age are determined by the POS tags. Research suggests that older users have a higher prevalence of 3rd person pronouns, determiners, adjectives, and conjunctions, while younger users tend to use more 1st person pronouns, nouns, interjections and adverbs (Brandt et al., 2020; Sloan et al., 2015). The prevalence of these POS tags is calculated for each comment relative to the total length which results in a value between -1 and 1 (-1 being a high prevalence of the old age indicator and 1 being high prevalence of the young age indicator). In order to calculate the grade level in the context of the UK, I adapted the Flesch-Kincaid grade level score, through a ratio of average sentence length and number of syllables per word (Chavkin, 1997) to the equivalent UK grade levels (see section 3.2.4 for details). The analysis was implemented for the Fireworks discussion on Facebook as it had the most comments out of my sample as well as a high average comment length.

Using the results from the grade scores and part-of-speech analysis, users are classified in various groups (e.g. low education, high education) using clustering analysis as described in section 3.2.4. The results of the clustering analysis are visualised in Figure 45. Following this, sentiment analysis and LDA topic modelling (Blei, Ng and Jordan, 2003) of the posts of users in these various groups will be conducted in order to better understand whether these various socio-demographic groups vary in the way they contribute to political discussions initiated by the UK Parliament.

Figure 45: Three K-means clusters plotted against socio-demographic indicators of UK grade levels and age indicators



Several clustering algorithms including k-means and hierarchical clustering (Kaufman and Rousseeuw, 2009) are applied to the data incorporating the prevalence of age indicators (identified through POS tags) in combination with the UK grade levels. Each clustering was run through several optimization measures to evaluate the most suitable number of clusters finding 3 clusters using the k-means algorithm was optimal. These can be categorized as: (1)

low prevalence of young indicators and lower grade levels, (2) high prevalence of young age and higher grade level, and (3) higher age and higher grade level.

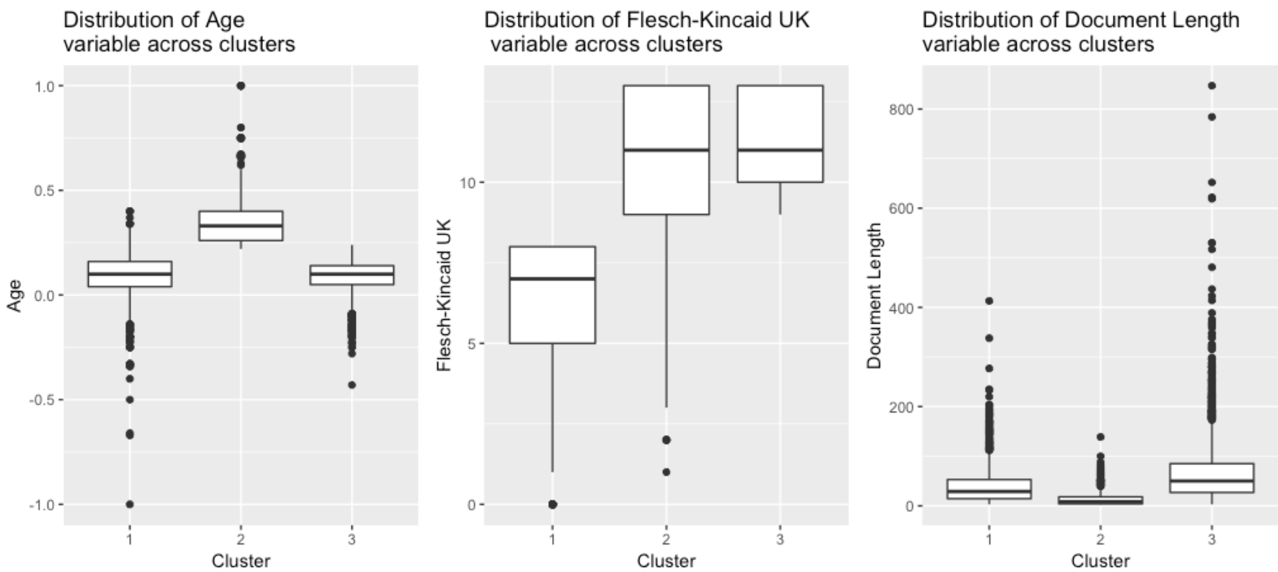
Table 6: Summary of indicator variables per cluster

Cluster	UK Grade level		Age indicator		Document length		Size
	Mean	SD	Mean	SD	Mean	SD	
1	6	2	0.096	0.102	39.3	36.2	2214
2	10	3	0.360	0.145	13.7	14	724
3	11	2	0.093	0.074	66.4	63.7	3074

With the socio-demographic clusters obtained, I can now descriptively analyse these clusters with respect to the socio-demographic indicators and textual and semantic features. In a second step I will also explore whether there are any differences in the expressed sentiments or topics between the clusters in order to better understand whether these various socio-demographic groups vary in the way they contribute to political discussions initiated by the UK Parliament. This analysis will be conducted in the next section 6.3 along with the general topic modelling analysis as well as in the section 6.4 along with the general sentiment analysis. For now, the focus will be on understanding who is represented in these three socio-demographic clusters.

Table 6 presents a summary of the descriptive statistics of indicators in each of the three clusters. The clusters can be grouped according to these variables. Regarding the UK Grade level, clusters (2) and (3) are both more aligned with secondary school education but are very different in terms of comment length, while cluster (1) is more aligned with primary education in the UK. However, when looking at the age indicator it appears clusters (1) and (3) are more aligned and have a similar prevalence of each indicator. This is also supported by the graph in Figure 45 which shows clusters (1) and (3) are very similar in age (y-axis) however differ greatly on the UK Grade level (x-axis). In general, the average positive age indicator across all clusters suggests the majority of participants in this discussion were of a young age. This could be a result of Facebook being a popular social media channel in the young bracket (Ofcom, 2019) shown by the vocabulary and syntax used.

Figure 46: Distribution of socio-demographic indicator variables (age, Flesch-Kincid UK) and ducument length across clusters identified through k-means algorithm



While the values in Table 6 give the means and standard deviations for each variable, the boxplots in Figure 46 shows a more detailed distribution. These show that generally cluster (2) has the largest interquartile range in the four variables but this can be down to the smaller size of this cluster causing more variability. Nevertheless, while cluster (3) has the highest grade level at 11, it also appears to have a lot of outliers as and a large range in document length. This contrasts with cluster (2)'s similar high average grade level but a much smaller variability in terms of document length. An analysis of the topics discussed in these clusters is given in the next section 6.3. While this analysis provides some insight into the demographic distribution of users, it is not possible to validate these results because the profile information of the users is not available. Therefore, a limitation of this analysis is that it is only speculative.

6.3 What do they find important? – Extracted LDA topics

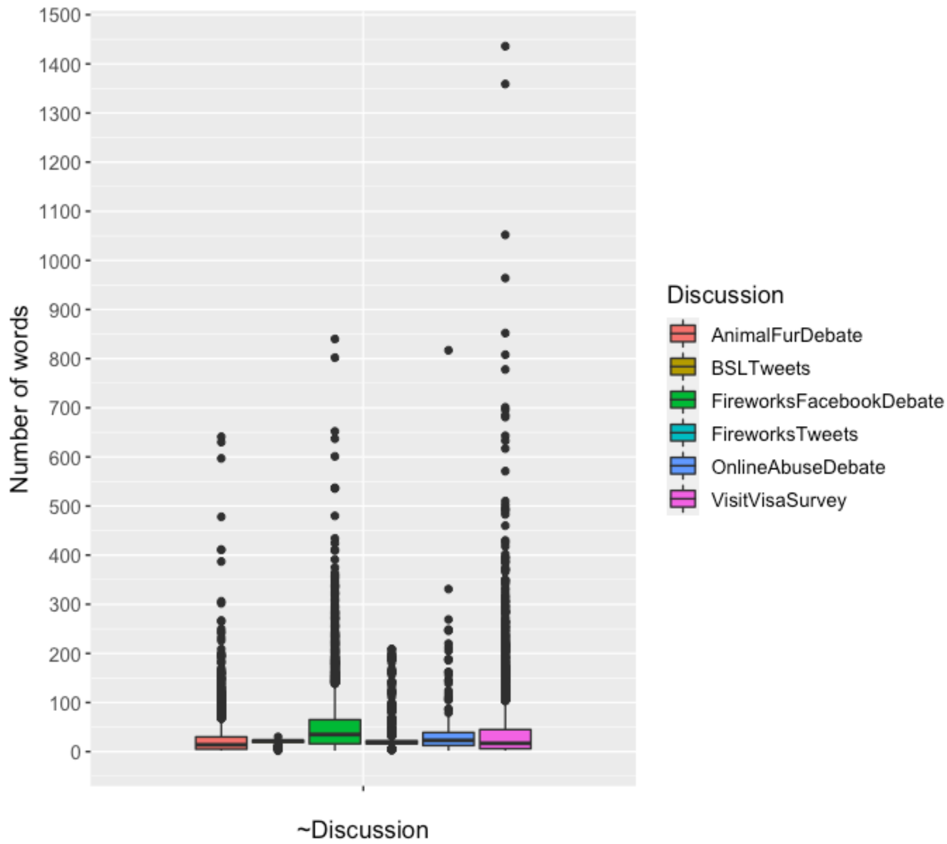
So far, this chapter has explored various aspects of five different engagement activities. This has concentrated on exploring the backgrounds of the participants through analysis of when they were active online and how they interacted with each other and through the linguistic characteristics of their posts. I found that the debates differ in terms of the pattern of submissions during the course of a day and during the lifetime of the debate. Some discussions follow a normal working pattern of 9am to 5pm while others indicate participants submit comments throughout the day. I have also found that the network of Twitter interactions differs heavily between two discussions, Fireworks and British Sign Language (BSL), partly due to the use of either a general or bespoke hashtag and the varied subject matter.

Now that I have gained some understanding of the types of people engaging in these digital discussions initiated by the House of Commons, I can delve deeper to understand what within the discussion topics is most important to them and what the participants in these debates are actually arguing for. However, one challenging aspect of comparing these debates is the varied subject matter which is covered by Parliament's digital engagement activities. The bigram networks shown in section 6.1 provide an overview for each discussion, but this is based solely on word frequencies. Increasingly, topic models, a machine learning approach (see section 3.2.3) are used to uncover the primary topics in a particular textual dataset (Yan *et al.*, 2013; Blei, Ng and Jordan, 2003). In this section, I use topic models to explore the range of issues raised in each discussion.

As explained in section 3.2.1 textual data needs to be first cleaned and pre-processed, before topic models can be implemented. Following the pre-processing of the data, I firstly examined how substantial each comment is, based on its length by calculating the row totals for the document-term matrix. Focussing on the distribution of comment lengths, Figure 47 shows that the Visit Visa survey has the largest total range of comment lengths with the largest over 1400 words, as well as the second highest average comment length of 43 words This is expected as a survey does not place as many word restrictions on submissions as for example as Twitter does. On the other hand, as this discussion consisted of individual submissions which could not see other participants' comments, participants may have felt they had more license to give longer and more personal information.

Fireworks Facebook debate and Online Abuse (also a Facebook Debate) had the largest (50 words) and third largest (43 words) average comment lengths respectively, but I can see that the interquartile range is much larger for Fireworks comments on Facebook. All but one of Online Abuse comments were under 350 words, with the exception being a particularly long comment of over 800 words. The Fireworks discussion on Twitter has a small range of comment lengths, but still has some comments just over 200 words. The BSL tweets had a very small distribution of words per comments as well as a much smaller range than the other discussions. The Digital Engagement team also do not use Twitter for the type of engagement where they are generally seeking views and participation form the public, therefore to emulate the team's current practices, I will perform topic modelling on only the Facebook discussions and survey. Looking at the comment lengths of the text allows me to make comparisons irrespective of subject matter and places more emphasis on the engagement platform used.

Figure 47: Distribution of number of words per comment in digital debates



The Latent Dirichlet Allocation (LDA) topic model is a computational method for identifying topics in textual data (discussed in section 3.2.3), several topics are proposed. Training the model on a sample of all the discussions resulted in very weak topics which could not be well interpreted. This led to each discussion mainly being categorised into one single topic. As my aim is to explore the sub-topics within each discussion, separate models are used for each. The number of topics chosen for each debate differs depending on the optimum topic number as defined by LDA Tuning package in R (Nikita, 2016), which is described in section 3.2.3. This is determined by a maximisation of the Griffiths2004 metric and a minimisation of the CaoJuan2009 and Arun2010 metrics. I set the package to evaluate topic numbers between 2 and 10 as the Digital Engagement felt that anything over 10 would be unhelpful for them to meaningfully analyse, and they had not encountered any previous digital discussions with very high numbers of topics (manually evaluated by them). I then manually examined the word distributions for the most likely optimum topic number for each discussion to see which led to the most useful results.

Table 7: Fireworks Facebook Debate - LDA statistical validation metrics

topics	Griffiths2004	CaoJuan2009	Arun2010
10	-811419.4*	0.103	1876.46*
9	-828116.4	0.126	1932.99

8	-822495.5	0.099*	1971.56
7	-847619.8	0.109	2026.36
6	-842654.7	0.100	2087.21
5	-864158.5	0.133	2180.43
4	-880631.4	0.111	2268.49
3	**-902843**	**0.128**	**2410.51**
2	-931106.8	0.192	2611.72

Table 8: Animal Fur Debate - LDA statistical validation metrics

topics	Griffiths2004	CaoJuan2009	Arun2010
10	-179682.9*	0.126	415.18*
9	-180899.4	0.102	419.16
8	-182126.0	0.104	433.50
7	-183988.0	0.105	447.76
6	-186526.6	0.142	470.81
5	**-188990.5**	**0.083**	**486.41**
4	-192962.8	0.136	515.31
3	-197703.4	0.019*	539.35
2	-203817.5	0.222	609.75

Table 9: Online Abuse debate - LDA statistical validation metrics

topics	Griffiths2004	CaoJuan2009	Arun2010
10	-38125.58*	0.100	44.1*
9	-38141.87	0.104	46.5
8	-38056.80	0.091	48.13
7	-38358.22	0.088	50.73
6	**-38445.95**	**0.071**	**53.2**
5	-38854.11	0.092	58.42
4	-39399.92	0.119	63.16
3	-40211.15	0.122	69.9
2	-41618.10	0.042*	77.97

Table 10: Visit Visa Survey - LDA statistical validation metrics

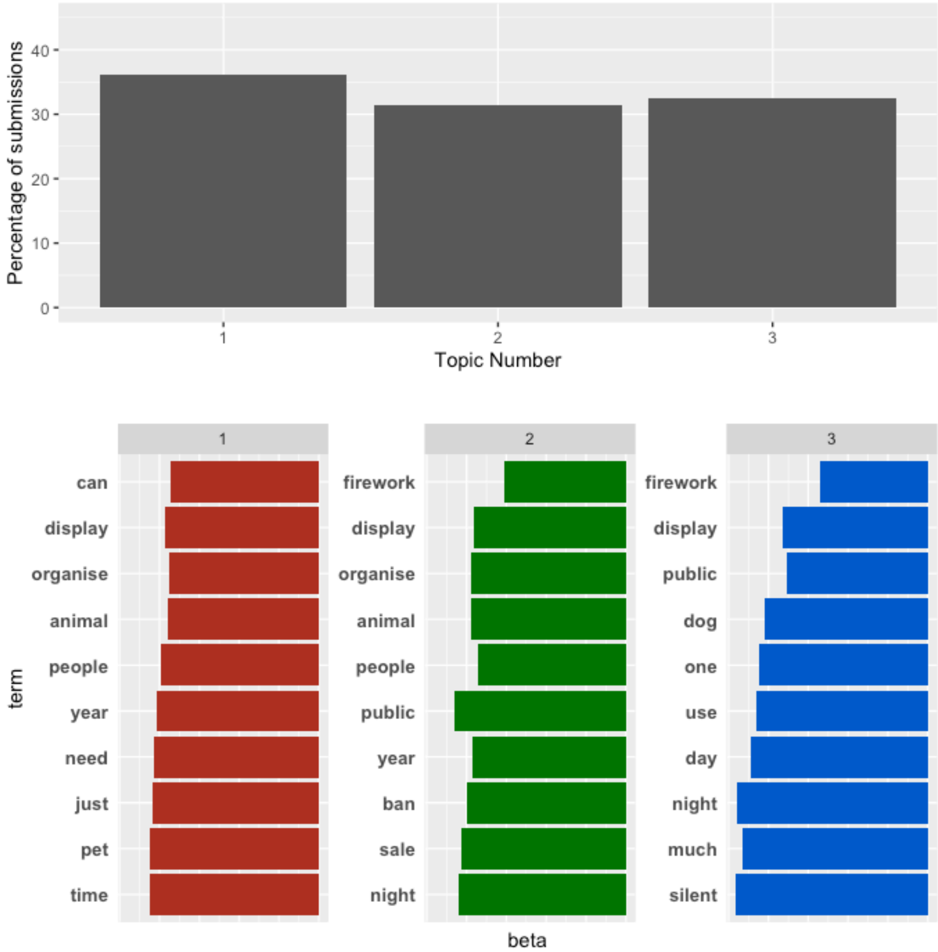
topics	Griffiths2004	CaoJuan2009	Arun2010
10	-434141.6	0.077	481.32*
9	**-433019.0***	**0.061***	**490.98**
8	-437517.7	0.078	511.03
7	-442146.8	0.089	535.79
6	-447296.8	0.069	551.44
5	-454622.3	0.075	582.14
4	-465323.1	0.155	620.60
3	-475214.2	0.148	660.30
2	-495380.2	0.072	714.21

The optimum topic numbers according to the metrics are three for Fireworks, five for Animal Fur, six for Online Abuse, and nine for Visit Visa shown in Table 7, Table 8, Table 9, and Table 10 respectively. These metrics show that the Griffiths2004 and Arun2010 measures almost consistently optimise the highest number of topics, while the CaoJuan2009 measure has more variation in the topic number optimised in each discussion. However, when I examine the word distributions in these optimum topics manually, I find that there is often not a specific separation of topics which would be helpful for the Digital Engagement team to interpret. I therefore, chose the topic number which I believe produces the most interpretable results with topics which were specific and distinct enough from each other to be meaningful. Therefore, while topic modelling is a very useful algorithm for textual data, choosing the correct number of topics when setting the parameters of the model can be down to simple trial and error when using optimisation techniques. This subsection will focus on the alpha and beta values of the topic model which provides information about how the documents were distributed among the topics, and how the words were distributed. The alpha scores for the Animal Fur, Online Abuse, and Visit Visa discussions were all low values (under 1.0) and suggest there each comment is represented by only a few topics. This implies each comment is quite different to the next. On the other hand, the Fireworks discussion had a very high alpha value which suggests that the comments are more similar to each other.

The Animal Fur debate (Figure 50) and Visit Visa survey (Figure 52) are dominated by one single topic, 5 and 4 respectively. While the Fireworks debate (Figure 48) and the Online Abuse debate (Figure 51) are much more even in terms of the percentage of comments categorised into each of their topics. The beta scores (per-topic distribution) give the probability a term belongs to a topic (Blei, Ng and Jordan, 2003). Using this principle, the model assigns a particular word to a topic by finding the maximum beta score for that word across all topics. Ordering the beta scores in descending order gives the words with the highest probability of being in any of the topics.

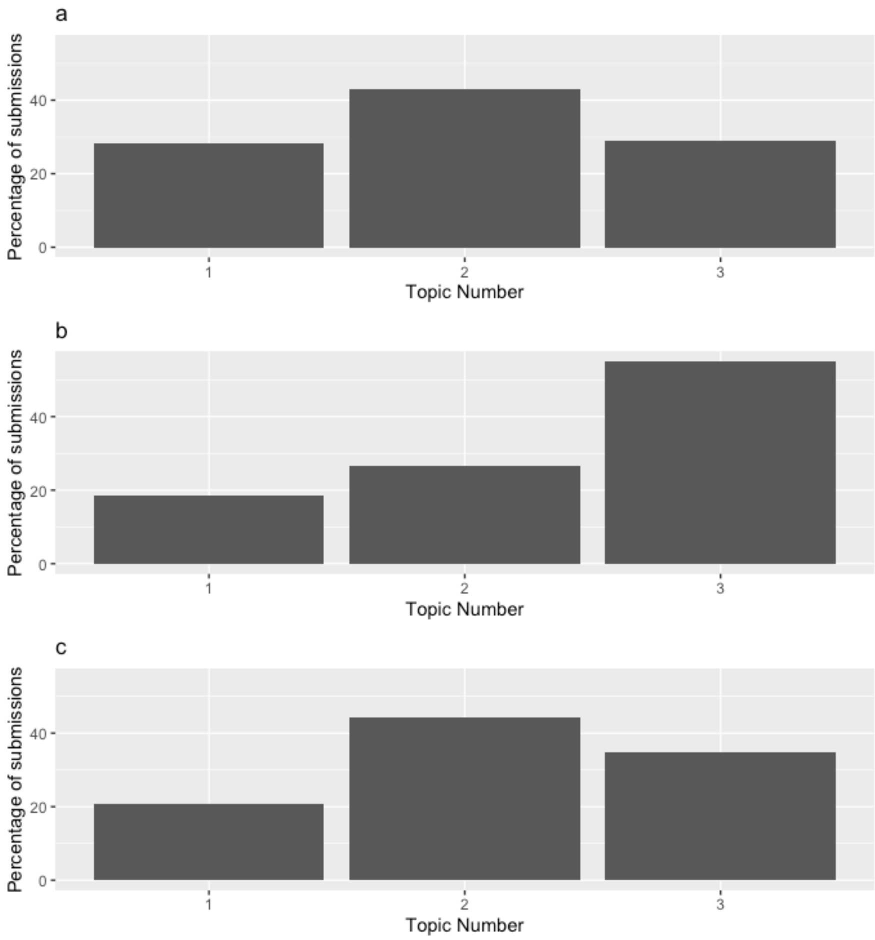
In the Fireworks debate, participants appear to discuss each topic equally with no single topic representing many more comments than others (Figure 48a). This suggests the discussion was varied and covered a range of three subtopics. This topic model has a high alpha value of 50.14 suggesting the comments are made up of a mixture of different topics and the comments are similar to each other. Figure 55b shows a selection of words in each topic ranked by the highest beta scores. While this gives some insights into the nature of the topics, many words are repeated across topics such as ‘display’ and ‘organise’. This shows a slightly different interpretation of the topics, particularly topics 1 and 2 which reference the effect of unregulated firework displays on military personnel with words such as “brigade” and “ShoulderToSoldier” which is a charity to support armed forces families (ShouldertoSoldier, 2020). This provides an insight into the types of users or organisations present in the discussion as well as their main areas of concern. Specifically, the emphasis on “organise” suggesting this is the type of firework display most participants would prefer, and “animal” showing that participants are particularly concerned with how fireworks are affecting their pets. While this LDA model does extract different topics, the difference between topic 1 and topic 2 in terms of the distribution of words in each appears very small. These two topics have a large overlap with only a few differences in the words and suggests the model has not performed very well in this discussion. Furthermore, the high alpha value of this model indicates the comments are similar to each other and may therefore be difficult to distinguish.

Figure 48: Fireworks Facebook Debate – Topic distribution (a) and selection of words per topic (b)



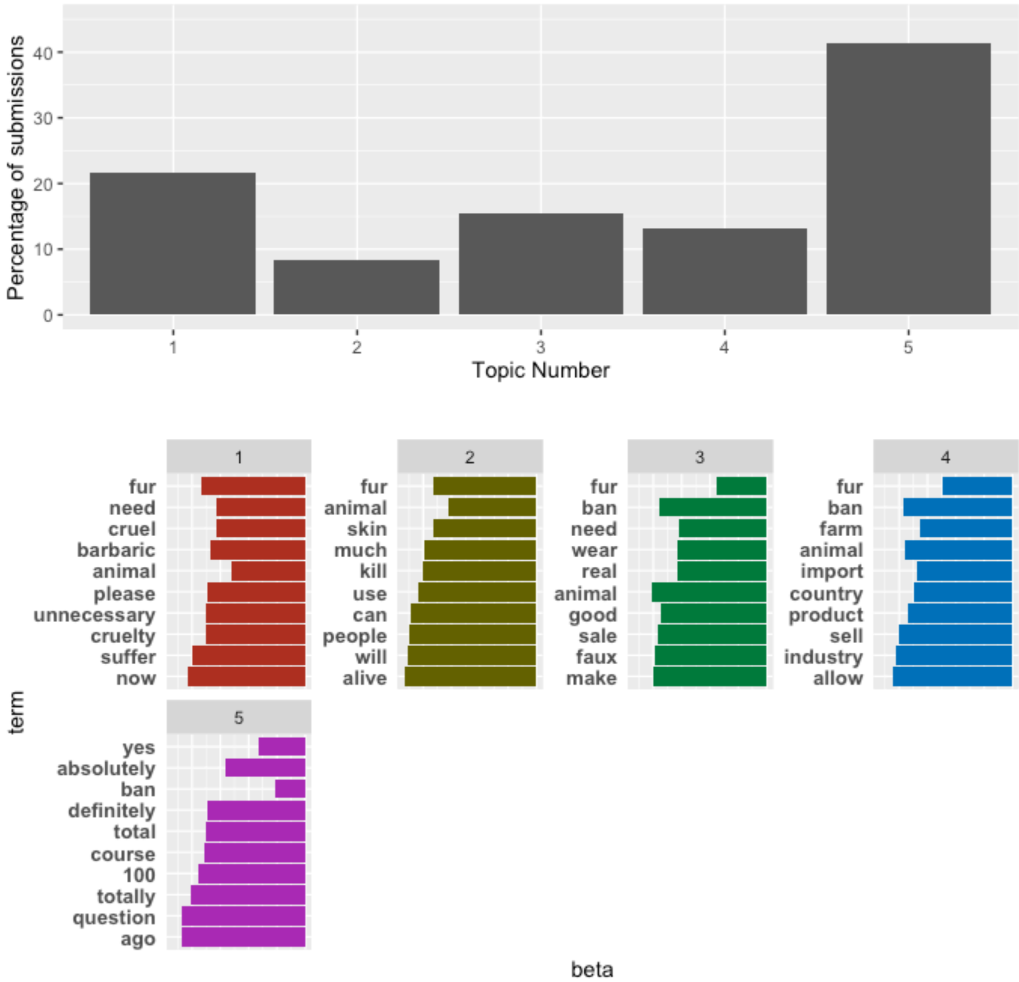
In section 6.2.3 I categorised the users based on their comments posted in Fireworks discussion into clusters that represent certain socio-demographic attributes. These clusters were summarised in Table 6. The question now is to what extent do the different socio-demographic clusters differ in what they contribute to the digital discussion. Having run the LDA topic model for this Fireworks discussion as shown in Figure 48, I can also understand the specific topics each cluster was most focussed on. There was also the option of running separate LDA models for each of the clusters to see what topics they each focussed on, however the purpose of this exercise is to explore how participants respond to online engagement sessions and the inclusion of the socio-demographic analysis is intended to complement this. Therefore, as they are about the same discussion, I have chosen to apply the same LDA model for the analysis of the socio-demographic clusters using the posterior probabilities of the model. Figure 49 shows that Cluster 1 which was characterised by low grade levels and cluster 3 characterised by higher grades both were slightly more interested in topic 2 which features words such as "animal", 'kennel", and "soldier". Cluster 2 characterised by younger age and higher grade level were considerably more interested topic 3 than any of the other topics. This topic covered issues such as using silent fireworks instead of conventional ones and also the punishment of those who have unlicensed displays. Therefore, by combining the clustering results with LDA topic modelling results I can see which topics are more important to different sections of society I have identified. The younger participants appear to be more interested in alternatives to fireworks while older participants irrespective of their educational background were more focussed on the issues fireworks present.

Figure 49: Fireworks Facebook Debate - Topic distributions per socio-demographic cluster 1 (a), cluster 2 (b), and cluster 3 (c)



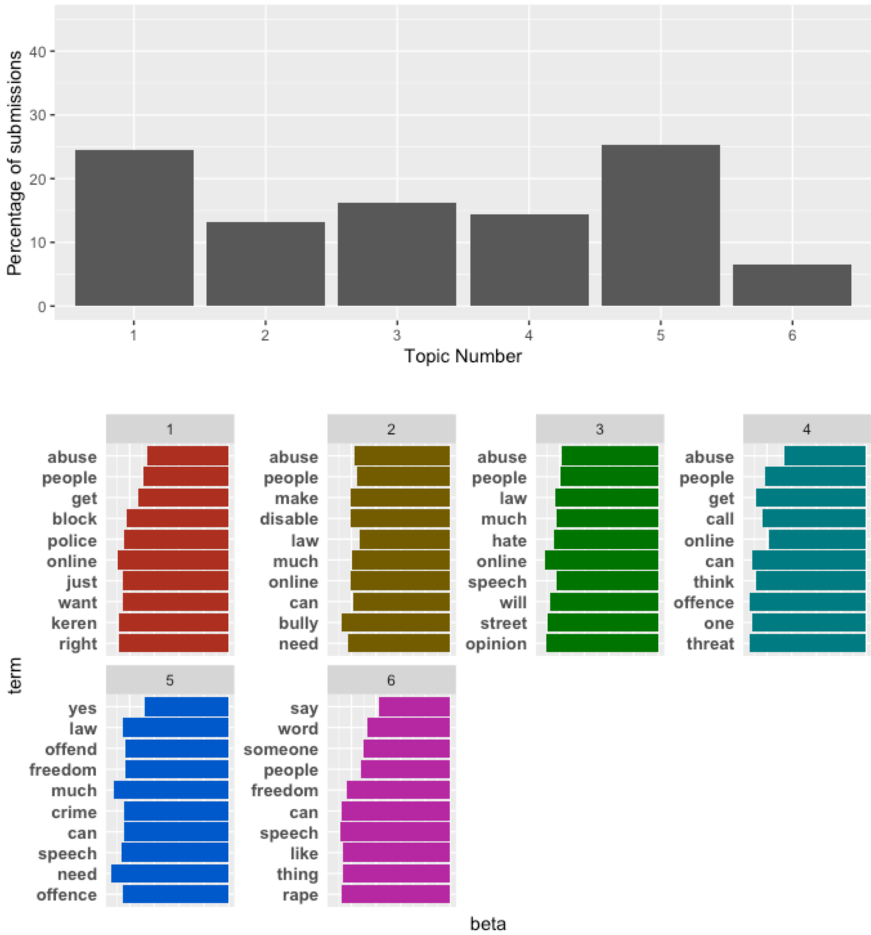
In the Animal Fur debate, the topics and distributions uncovered are displayed in Figure 50. This topic model has a low alpha value of 0.78 suggesting the comments are made up of a small number of topics and the comments are quite different to each other. They reveal that 42% of comments are categorised into a single topic (5) which is primarily concerned with agreeing the ban of the sale of animal fur, with words such as "'totally", absolutely", and "definitely. The topic containing the fewest comments is topic 2 which talks about how people kill the animals and prepare their skin and fur for sale. This topic has words such as "kill", "people", and "skin". Therefore, as with the bigram network (Figure 31), it appears participants in this debate can be categorised as being either more focussed on the specific details of fur production, or on the emotional and moral implications of selling animal fur. Therefore, one could conclude that the majority of participants were indeed on-topic and talking about the impact of selling fur in the UK, but used very emotive language to convey their views. The Digital Engagement team can use these insights to inform MPs that they could concentrate on the emotional aspect of animal fur production as this is an important issue for the public.

Figure 50: Animal Fur Debate - Topic distribution (a) and selection of words per topic (b)



In the Online Abuse debate, Figure 51 reveals that there are two topics, 1 and 5, which contain 50% of all comments. These topics are primarily concerned with involving the police in the abuse of disabled people online, and the need to use the law. This topic model has a low alpha value of 0.083 suggesting the comments are made up of a mixture of only a few topics and the comments are rather different to each other. Topic 6 had the least number of comments where participants discussed other types of crime that had been threatened online such as "rape", "murder" and "attack". This shows that while the majority of the discussion was focussed on the process of involving the police and the law in cases of online abuse of those with disabilities, there was also some smaller discussion about more serious threats.

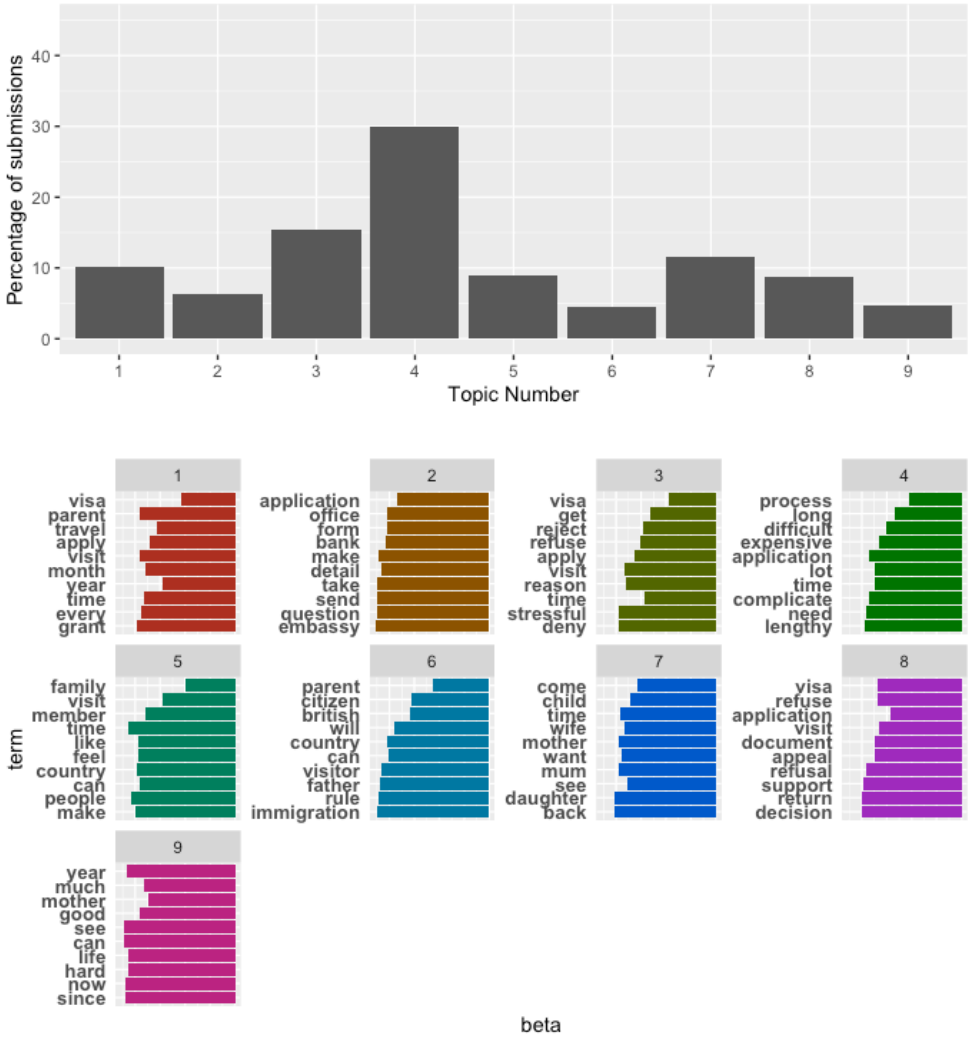
Figure 51: Online Abuse Debate - Topic distribution (a) and selection of words per topic (b)



The Visit Visa survey featured some of the longest comments out of all of the discussion and also covered a varied range of issues (Figure 52). Similarly, to the Animal Fur and Online Abuse discussions, this topic model has a low alpha value of 0.87 showing the comments are made up of a small number of topics and the comments are quite different to each other. The most popular topic was number 4 with 30% of comments shows that this relates to the length and cost of the application process and difficulties users found completing the documents. Other areas of interest in this discussion was the stress incurred by applications being rejected and the different reasons given by the home office for the visa denials in topic 3. Mentions of different family members being unable to visit their relatives in the UK are also prominent in topics 5 and 7. Topic 9 is the only topic which does not appear to have any particular concrete themes such as those that have been extracted from the other topics.

This is one difficulty of using topic models such as LDA, where the results are algorithmically produced but require a human to make sense of the topic results which may not always yield interpretable topics. Nevertheless, when analysing hundreds or thousands of comments in a short space of time, as is done in the UK Parliament, any automation uncovering some core themes in a discussion will ease resources of staff time. Furthermore, the summaries of topics, temporal activity and sentiment analysis (as will be shown in the next section) successfully emulate the engagement activity summary reports that the Digital Engagement team currently produce for MPs. In doing so, the team is still able to use this analysis to produce the same information and insights into a discussion that they would have previously completed manually. In the case of Twitter discussions, they are also able to take advantage of the social network analysis to observe how participants interact with each other and the effect this can have on discussions.

Figure 52: Visa Debate - Topic distribution (a) and selection of words per topic (b)



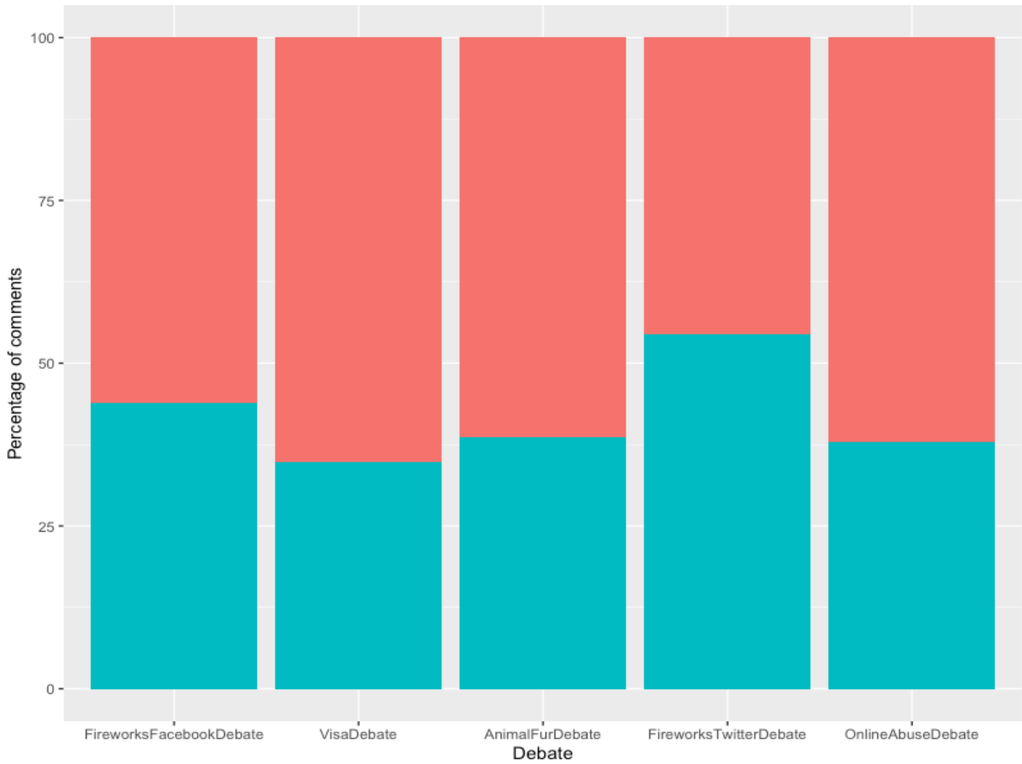
6.4 And how does that make them feel?

My final question in this section is exploring how the participants feel about the topics they are engaging with. The majority of people who participate in an online session, do so because they have a vested interest in the topic and it is an issue close to their hearts. This is less so because they are especially interested in engaging with the institution. As such, I can expect the language use to be relatively emotive due to participants' personal connection to most of the topics. In this section, I use three lexicons for sentiment analysis (see section 3.2.2. for details) to explore whether different lexicons, Bing, Afinn, and NRC, lead to different interpretations of the sentiments expressed in each discussion, and which reveals the most accurate picture of participants' sentiments.

6.4.1 Binary Classification

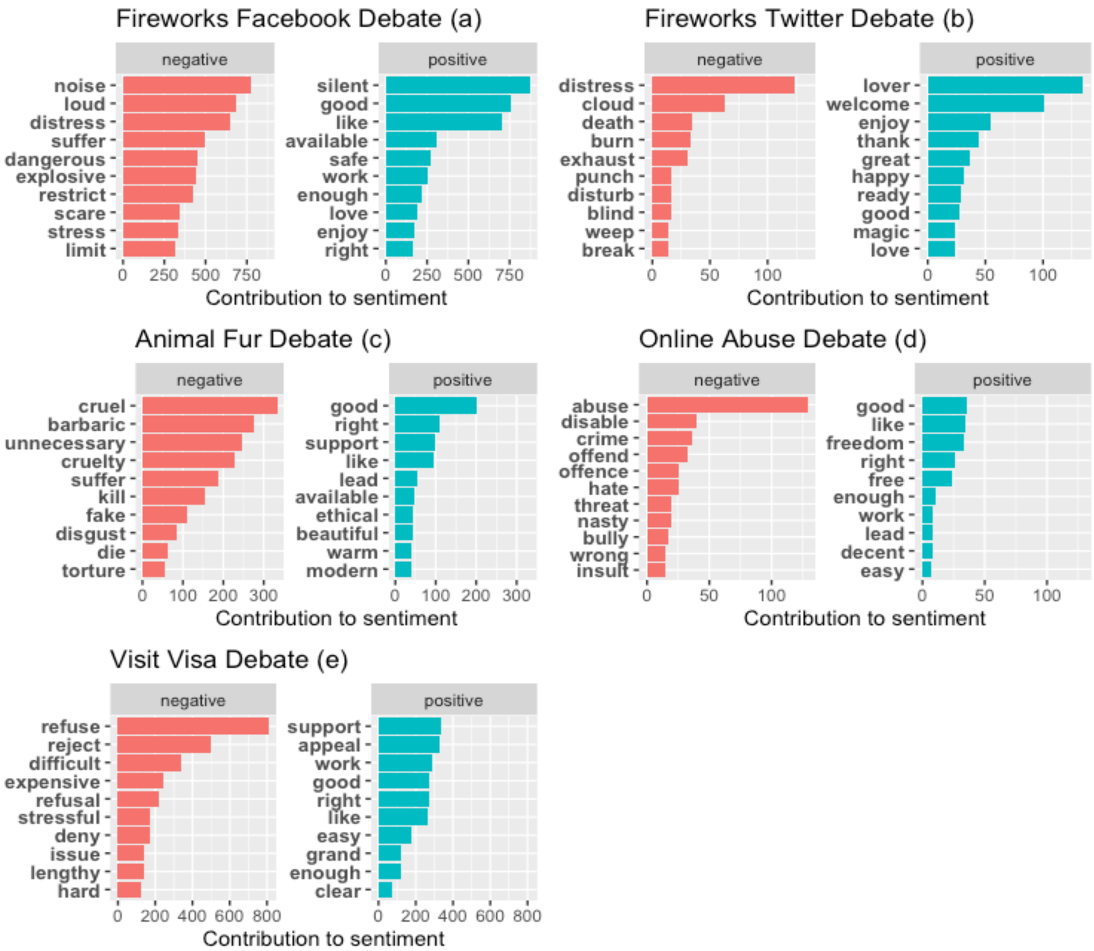
Applying sentiment analysis to the comments across the examined digital discussions provides a list of words and their probabilities of being positive or negative. I can calculate the proportions of comments in each debate belonging to either sentiment category. Starting with the Bing lexicon, this particular classification contains 6788 words already with an associated sentiment to decide how to categorise the new dataset (Liu, 2018). Figure 53 shows the percentages of comments classified as either positive or negative. It reveals that between 55%-65% of comments are negative in all debates with the exception of the Fireworks discussion on Twitter. This discussion was the most positive with 57% of the tweets belonging to the positive sentiment category. This is in contrast to the same debate on Facebook which follows the pattern of other discussions as being more negative than positive. This contrast is most likely due to the fact that the hashtag used in this discussion was a general one (see section 6.2.1) and thus included lots of comments that were not about the e-petition at all but rather about fireworks in general, which are often associated with positive events such as New Year celebrations.

Figure 53: Bing Lexicon – percentage of comments categorised into sentiments for digital debates (blue is positive and red is negative).



The examination of words contributing the most to each sentiment shows that the lexicon successfully categorises words based on their sentiment as a human would intuitively (Figure 54). Data from the Fireworks Facebook debate shows that the words with the most negative sentiment in the comments are ‘noise’, ‘loud’, ‘distress’, while positive sentiment are ‘silent’, ‘safe’, and ‘love’. However, the same debate on Twitter contains mostly very different words categorised to each sentiment with “distress, “cloud” and “death” as negative, and “lover”, “welcome” and “enjoy” as positive. The Animal Fur debate negative sentiment includes ‘cruel’, ‘barbaric’, and ‘unnecessary’, while the positive are ‘support’, ‘right’, and ‘ethical’. The Online Abuse debate overwhelmingly categorises the word “abuse” as negative, along with “disable” and “crime”, while positive words are “freedom”, and “right”. Finally, the Visit Visa discussion contains negative words such as "refuse", "reject", and "expensive", and positive words "support", "appeal", and "work". Results such as this are useful for the Digital Engagement team as they can begin to understand how the participants are expressing themselves with respect to the different discussions. Some of these words are only positive of negative in a particular context rather than examining them individually. So although this gives me some information, using unigrams or single words for sentiment analysis does not give a great deal of context.

Figure 54: Bing Lexicon – representative words per sentiment in each digital debate



Instead, I can focus on the comments themselves, however as the goal of these analyses is to quickly pinpoint the public's opinions, some pre-processing is first necessary. The method of calculating sentiment for a comment involves assigning a sentiment category to each word in the comment, and then finding the category with the most occurrences. If a comment has 3 positive words and 7 negative words, it is classified as negative. However, this method of classification can provide unreliable results as either extremely long comments or extremely short comments would take precedence when sorting comments in descending order of percentage or raw number of words. Therefore, the comments with the largest word count will feature highest on both sentiment scales.

To get around the problem of raw comment length, a percentage or proportion of the total words in the comment is used to order the sentiment scales. In this case, when ordered, some comments receive 100% positive or 100% negative scores because they have very short word lengths. The negative scores are accurate for the short messages i.e. "Hate them", "Without a doubt!". While these comments are indeed negative, their short lengths does not provide much context for the Digital Engagement team. So, including a threshold for very long and very short comments provides a compromise between interpretability and accuracy. Values for these thresholds can be decided by only including comments lying within the interquartile range. Therefore, if I set the threshold to this range, I can see the sentiment for the most representative of the comments posted.

Table 11 displays the specific comments the lexicon has identified as being the most positive and the most negative in the dataset. In the positive category, the Bing lexicon successfully manages to extract positive comments for the Fireworks on Facebook and Twitter, Animal Fur, and Online Abuse debates, but the Visit Visa survey comment is not a very positive one. This discussion had the highest percentage of negative comments among the dataset with 63% and could highlight a potential disadvantage of the lexicon. When the overall distribution of comments in a discussion is highly negative the lexicon struggles with identifying positive comments. On the other hand, the Bing lexicon performs well in the categorisation of negative comments across four discussions with the exception of the Fireworks discussion on Twitter, where a comment was categorised as the most negative, which however was not necessarily negative but did include the words "death", "hate" and "punch" referring to the metal band Fire Finger Death Punch. This example shows the Digital Engagement team that a lack of specificity in the organisation of the online engagement session (in this case, the generic hashtag) can have consequences that reach through multiples areas of analysis, and causing the results to be unreflective of the intended discussion.

Table 11: Bing Lexicon - most representative comments per debate

POSITIVE	
FIREWORKS (FB)	*"Nooooo. These are fun and exciting. I only live once so grab every bit of enjoyment you can and considerate fun with fireworks has brought lots of lovely memories for my friends and family."*
FIREWORKS (TW)	*RT @McZameth: Clear skies for the start of the amazing #fireworks. Last photo I promise lol. #AustraliaDay #Perth #LoveMyCity @cityofperth*
ANIMAL FUR	*"It works for diamonds and fair trade goods. It also works in places like Australia"*
ONLINE ABUSE	"*Thank yo for your intelligent and researched comments folks....memo to self...check first. I now agree...our Fascist government has enough powers!"*
VISIT VISA	"*Reasonable process. But some questions on application not very clear. Refused visa but no appeal. Not clear refusal notice."*
NEGATIVE	
FIREWORKS (FB)	*"Fireworks in an uncontrolled environment create fear and anguish in humans and animals, cause problems for the emergency services and can cause death, destruction and injury..."*
FIREWORKS (TW)	*RT @AwesomeMetal: #NowPlaying #Five Finger Death Punch - Falling In Hate - https://t.co/rxU9mWD2AV #BOA17 #HappyNewYear #Brexit #Fireworks*
ANIMAL FUR	*"Is this a rhetorical question?? Cruel and completely unnecessary suffering to animals simply for vain humans."*
ONLINE ABUSE	"*Nothing can be as bad as the number of people they've killed with their cruel, heartless policies! Isn't that classed as abuse?!!"*
VISIT VISA	"*it's expensive, chaotic, and very tiring for ailing parents."*

6.4.2 Non-Binary Classification

The analysis above was completed using the Bing sentiment lexicon, a general lexicon that categorises negative words and comments well, but appears to struggle with some positive comments. Another alternative is to use a lexicon specific to Twitter or social media data. The NRC Word-Emotion Association Lexicon (aka EmoLex) uses a wider range of sentiments and emotions including, "anger", "anticipation", "disgust", "fear", "joy", "sadness", "surprise", or "trust" as well as "positive" and "negative". There are 14,182 words in total conveying around 25000 senses (as one word can be categorised as multiple emotions) and was manually created using Amazon Mechanical Turk (Mohammad and Turney, 2010; Mohammad and Turney, 2013).

Not all of the categories will be intuitive or even useful for every debate, so it requires a human eye to manually evaluate the results. In my analysis, the positive and negative categories have been removed as I have explored them in the previous chapter and I will solely focus on the specific sentiment categories, *anger, anticipation, disgust, fear, joy, sadness, surprise* and *trust.* Figure 55 displays the results for all the digital discussions cases as well as the mix of sentiments in each debate. Visit Visa has the next highest proportion of *anticipation* and *sadness* comments, while Animal Fur includes the most *disgust* and *fear* type of comments. Online Abuse discussion had the highest overall sentiment of *fear* followed by *anger*, while the Fireworks Twitter discussion featured the most number *joy, trust*, and *surprise* comments out of all five discussions.

Table 12 shows the largest single specific sentiments that categorise each debate and an exemplary comment for this sentiment from the respective debate. As with the binary classification, the comments all fall within the interquartile range in terms of comment length. As this NRC lexicon allows for words and comments to be categorised into multiple sentiments, it is up to the reader to interpret which sentiment category the comment is most suited to depending on the context of the discussion. For example, in the Online Abuse discussion, the comment categorised in the *fear* sentiment also featured highly in the *disgust* sentiment category. This comment mentions "threat" and "harassment" and is clearly coming from a frustrated user who is concerned about the online abuse of disabled people. Likewise, for the Animal Fur discussion the *fear* comment was also categorised as *sadness* and *disgust* which I would interpret it more closely as. This comment mentions the "pain, misery, horror, and suffering" of animals as a result of the fur industry and is an example of the emotive language used by participants in this discussion. Therefore, while the comment could be categorised as expressing fear of the future of animals involved in the practice, it is also clearly expressing sadness. The other main sentiment in the Animal Fur discussion is *disgust,* and an example of a comment in this category is in Table 12. This comment is well categorised and includes words such as "disgusting", "cruel", and "barbaric" which all convey the emotion of disgust well.

The Fireworks Facebook discussion also followed this pattern of comments with multiple categories with the *anticipation* comment also categorised as *trust*, while the *fear* comment was also categorised as *anger* and *sadness*. It appears there is an overlap with many comments categorised as *fear* also falling into the *sadness* category. The NRC lexicon does better at categorising the comments from the Visit Visa survey, compared to the Bing lexicon, with each of the three main sentiments (*anticipation* about a long wait for visa confirmation and uncertainty, *sadness* due to unfairness and discrimination, and *trust* in terms of lack of confidence in the UK) all returning relevant comments. The Fireworks Twitter discussion also correctly returns *anticipation* and *joy* for a tweet celebrating Eid. However, this lexicon struggles slightly with identifying the sentiment in the Fireworks Facebook discussion, categorising a comment about regulating fireworks as *anticipation* when it is in fact rather neutral. This may also be a difficulty where comments which do not clearly express any particular emotion are categorised into sentiments that may not make sense. This could be because the final sentiment category for a comment is decided based on the sentiment with the highest proportion of words in that comment. Therefore, some comments which only have a few words recognised by the sentiment lexicon may be incorrectly categorised, as appears to have been the case with the Fireworks Facebook debate *anticipation* comment.

Figure 55: NRC Lexicon - percentage of comments categorised into sentiments for digital debates

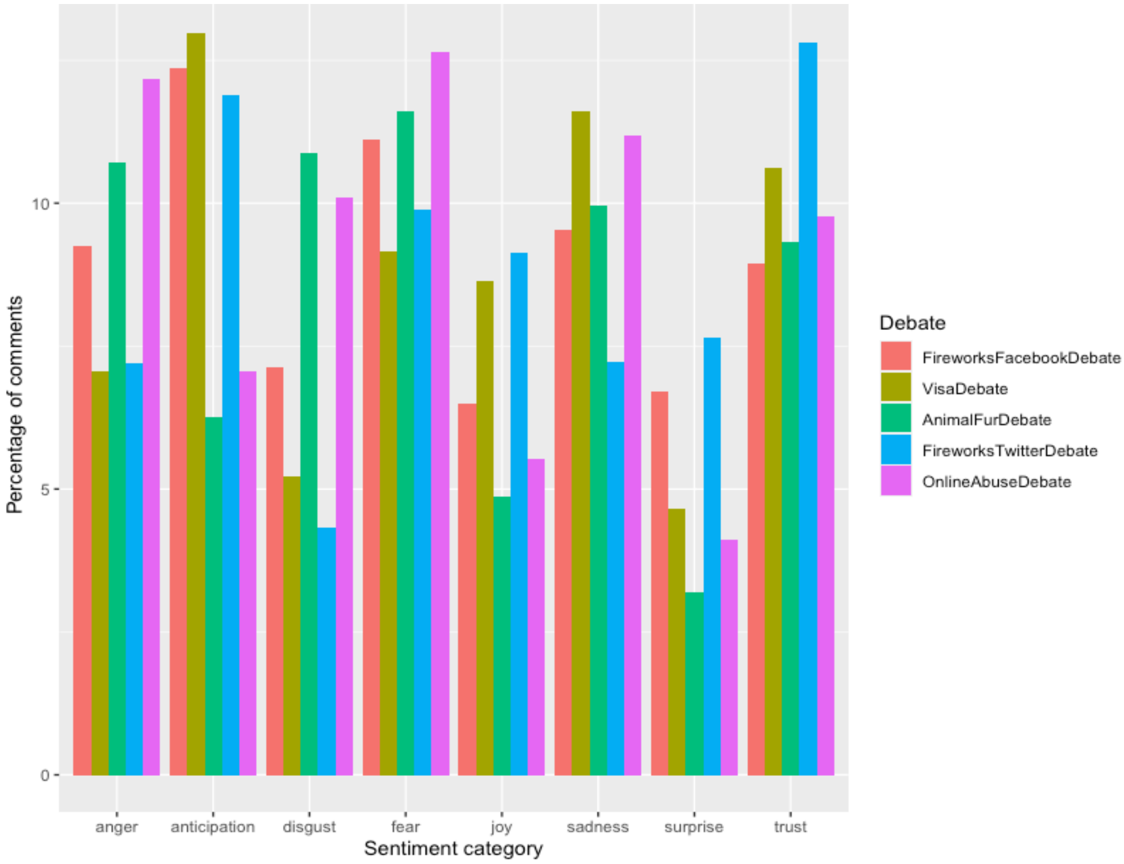


Table 12: NRC Lexicon - most representative sentiments and comments per debate

NRC WORD-EMOTION ASSOCIATION LEXICON

DEBATE	Top Emotion(s)	Example Comment
FIREWORKS (FB)	Anticipation	*"The use of fireworks definitely needs to be more regulated. They cause huge distress to people & animals alike. Should only be available to organised displays via a permit agreement at certain traditional times of year."*
	Fear	*"Fireworks should only be handled by professional people. They are far more dangerous now than they ever were and can now be considered lethal weapons. They have devastating effects on wild life, domestic horses and*

		domestic pets. All fireworks should be only available at organised events. There are also people who are suffering ptds or other mental health problems and autistic people who find the frightening. For everybody's sake restrict them to professionals only."
FIREWORKS (TW)	Anticipation	*RT @Shiven9876: #10WeekStreak @KHDA @IndianHigh_DXB @eidcelebrations watching&enjoying #EidalAdha2017 #fireworks #EidMubarak wid luv,peace*
ANIMAL FUR	Fear	*"Yes. It absolutely should be banned. It's completely unnecessary and its it is most certainly not worth the life, the pain, misery, horror and the suffering of another creature."*
	Disgust	*"I agree with the majority. Ban it, wearing another animals flesh is barbaric, cruel and disgusting. It's well past time to end such a horrific trade."*
ONILINE ABUSE	Fear	*"It already is, a threat is a threat however it's expressed, stalking me stalking, harassment is harassment ..... Censorship IS also a crime, a breach if human rights, but no one bothers to enforce it."*
VISIT VISA	Anticipation	*"A long wait, uncertainty about the visa meant I couldn't plan things in advance"*
	Sadness	*"Feeling of unfairness. Me and my wife pay more than 60 k in year in taxes. In return, there is a feeling of discrimination ..."*
	Trust	*"My mother was rejected and from that time I feel that UK is not fully my country, as my mother is my most important thing in my life"*

Both Bing and NRC categorise based on sentiment categories be they binary or non-binary, on the other hand, the Afinn lexicon uses a categorical valence scoring method. This ranges from -5 being the most negative to +5 being the most positive (Nielsen, 2011a). As with the other lexicons, the comment distribution of sentiments is plotted showing once again that the majority of documents have a negative sentiment. Figure 56 corroborates the findings of the Bing lexicon and shows Animal Fur in particular is a very negative debate, but also includes the largest range of sentiments with comments falling into the -4 and +5 categories, besides the Fireworks Twitter Debate which also features this range. All other debates have a range between -4 and +4. The majority of comments in the Animal Fur debate are in the -2 category which can be interpreted as negative, and only one word is respectively allocated to the most extreme categories of -4 ("bitch") and 5 ("outstanding"). The Firework discussion on Twitter also shows use of extreme language with values of -4 (also "bitch) and +5 ("breathtaking"). So it seems overall the range of emotions is rather comparable across the various discussions. The distribution of comments across the Afinn sentiment categories closely mirrors the results found by the lexicon's creator (Nielsen, 2011b, Figure 1). Specific comments within the comment length interquartile range are examined in Table 13 to test that the lexicon is performing correctly. As with the NRC lexicon, the values in Table 13 were chosen because they had the most number of comments categorised into them, so represented the prevailing sentiments of the discussion. All discussions are primarily negative with either -2 or -3 being the most popular sentiments, apart from the Fireworks discussion on Twitter which has a mainly positive sentiment of +2. This follows the interpretation of the Bing sentiment lexicon which also categorised the Twitter discussion as positive (Figure 53), most likely due to the lack of specificity of the hashtag. The Digital Engagement team can use this lexicon with issues where they are less focussed on finding the distinct emotions people are expressing as with NRC, but instead want to explore how extreme a discussion has been in terms of its comments' divergence from the -1 and +1 Afinn sentiment categories. This lexicon can act as an extension to the binary Bing lexicon with the negative Afinn values being a scaled representation of negative Bing sentiment and positive Afinn values representing different levels of the positive Bing sentiment. Therefore, the value for the collaborators in using this lexicon lies in its ability to categorise words and comments according to their valence and also the possibility of sub-setting a discussion based on the more moderate comments (i.e. those categorised between -1 and +1, or -2 and +2).

Figure 56: Afinn Lexicon - percentage of comments categorised into sentiments for digital debates

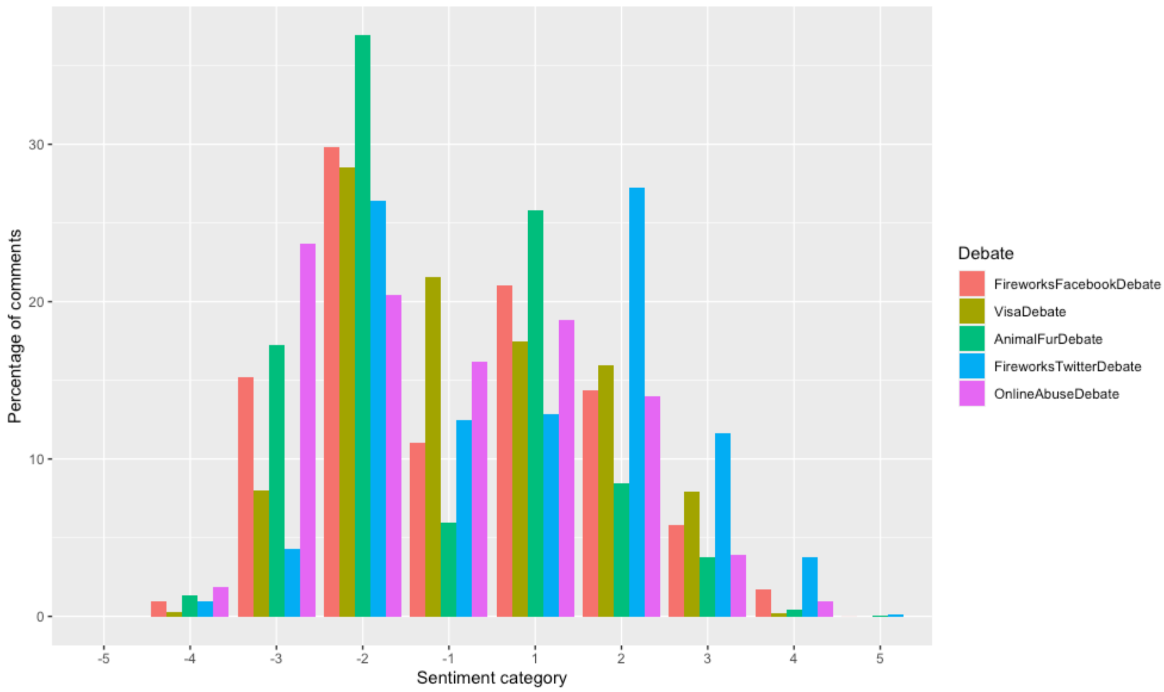


Table 13: Afinn Lexicon - most representative sentiments and comments per debate

AFINN SENTIMENT LEXICON

DEBATE	Top Emotion Score	Example Comment
FIREWORKS (FB)	-2	*"Please ban these fireworks my dog hates them it is so sad seeing how upset she gets....people let them off for months at all times of night with no regards to other people....don't forget it's not just animals that hate fireworks some people can find them as frightening"*
FIREWORKS (TW)	+2	*RT @Shiven9876: #10WeekStreak @KHDA @IndianHigh_DXB @eidcelebrations watching&enjoying #EidalAdha2017 #fireworks #EidMubarak wid luv,peace*

ANIMAL FUR	-2	*""Ban! Ban! Ban! No animal deserves to suffer!"*
ONILINE ABUSE	-3	*"Nothing can be as bad as the number of people they've killed with their cruel, heartless policies! Isn't that classed as abuse?!!"*
VISIT VISA	-2	*"it was approved, but was stressed, once was refused for silly reason a Document was missing, went for appeal which granted, so more was wasted"*

These three different sentiment lexicons while broadly showing similar results, reveal different characteristics of each debate. The Bing lexicon gives a positive or negative breakdown of the words and comments in each debate. This allows a basic overview of the sentiments and largely supports the more detailed categorisations of the other lexicons. The NRC lexicon allows for a much more granulated look at the data, revealing more specific emotions such as *anger*, *fear*, and *joy*. More detail brings a different perspective to the binary classifications and uncovers certain emotions that may be more valuable in the analysis of a particular debate. For example, where Bing categorised the Animal Fur debate comments largely as negative, NRC revealed there were more feelings of *anger* and *disgust* at the practice of killing animal for their fur and selling it. However, the increased detail of the NRC lexicon also leaves more room for mis-classification where some comments or whole emotions may not be useful for analysis. For example, the comment about regulating fireworks from the Fireworks debate on Facebook is categorised as *anticipation*, but appears to be more closely linked to *sadness* or even a neutral classification (which NRC does not have). Likewise, many of the comments in the other debates were categorised into multiple sentiment categories. Finally, the Afinn lexicon separates Bing's binary positive/negative system into a 10-step scale, allowing me to analyse the range of degree of sentiment in each of the debates. This lexicon provides valency scores which allow me to analyse the intensity of the positive or negative sentiments in the comments.

One feature all lexicons have in common is that the distribution of words differs from the distribution of comments. In all lexicons, the word distributions were much more even among the different debates, but the comment distributions were not. This suggests that overall, the types of language used in different debates is generally the same, but it is the combinations of the words which truly determine the overall sentiment and the differences between the debates. For this reason, focussing on the entire comments is vital to a clear interpretation of the data. This section has also included comparisons between the Fireworks discussion on Facebook and Twitter. The Afinn and Bing lexicons revealed that the Twitter discussion is generally much more positive than on Facebook. In fact, the Fireworks Twitter discussion is the only debate in the dataset which contains more positive comments than negative. This is not likely a result of the topic, because the same discussion on Facebook was predominantly negative, but more so a result of a non-specific hashtag used at the beginning of the engagement activity. This also reinforce how the setting up of an activity can have consequences on the discussion and how it will be evaluated.

6.5 Discourse platform and demonstration test set-up

The previous section has detailed the analysis of different discussions on a range of different online channels, Facebook, survey, and Twitter. These channels are used readily by the UK Parliament to engage the public online specifically with the intention of gathering their views and opinions on particular matters. The Fireworks discussion which was held on both Facebook and Twitter revealed differences in the discussion on each channel despite the topic of conversation being the same, albeit it is not clear to what extent the difference is due to the general hashtag used in the Twitter Firework debate, which resulted in many tweets that were not related to the e-petition. Nevertheless, the question of the effects of different communication channels on the types of responses participants are comfortable giving, remains, not least because earlier research has shown that participants on Facebook are more deliberative than on Twitter (Oz, Zheng and Chen, 2018), and local governments in Spain favour Facebook for a two-way conversation as compared to Twitter (Haro-de-Rosario, Sáez-Martín and del Carmen Caba-Pérez, 2018), Furthermore, as Twitter and Facebook are already used by Parliament, I can explore how a novel platform designed purposefully for online discussion influences the public, especially compared to the traditional channels that were not created with productive political digital discussions in mind, but rather have been repurposed to fit this need. Using existing channels has its advantages as there is already a ready-made community to participate on social media channels which other platforms may not have, but the demonstration tests I explain in the remainder of this chapter show that there are ways to use as purpose-built platform while still retaining the access to online communities.

Using a purpose-built platform can help to address the disadvantages of using social media platforms for online engagement such as the lack of moderation of comments, and difficulty in maintaining conversations among the participants. For example, on Facebook it can be difficult to maintain several conversations at once with the reply function and this feature may be off-putting to some participants specifically in discussions with many comments. The presence of moderation can on the one hand help to maintain civility in digital debates on the other hand it can also hinder engagement between Parliament and the public as participants may feel that they are not trusted by the institution to manage their own conversations. To overcome these various challenges, a range of different online platforms have been developed specifically for deliberative discussions online (Participatedb.com, 2019). These platforms have a multitude of features, each designed to improve or facilitate a specific aspect of online discussions. For example, Loomio[33] has a forum-like layout which allows users to keep track of conversations while also having different options of visibility for different discussions. This can be very useful if participants wish to discuss a particularly sensitive issue. OPIN[34] is comprised of six templates of participation exercises which can be customised depending on the type of engagement and the stage of the participation process. This platform was originally developed to engage the German youth in politics but can be used for all age groups. While Airesis[35] describes itself as "the social network fore-democracy" and allows users to create a proposal lasting between 1 and 30 days. This proposal can be shared to social media to invite friends into the debate. The platform provides information on how many participants are involved with each proposal and when it is ready to be voted on.

In this chapter, I will evaluate the use of the platform Discourse (Discourse.org, 2019) with three online engagement sessions held by three select committees in the UK Parliament (explained in detail in section 6.5.1). These discussions were held simultaneously with corresponding Twitter discussions in order to enable a comparison between modes of engagement in different platforms and whether these elicited different responses.

Increased engagement is one measure of success of these demonstration tests, however this project also aims to examine how engagement can be meaningful to the institution and contribute to policy-making. The success of these discussions can therefore be also evaluated on the basis of whether they have been included in published inquiry reports and questions to government ministers, showing that citizen input is meaningfully utilised and there is indeed a need for a new kind of discussion space in Parliament. However, the demonstration tests also revealed several limitations, some unique to the institution, which could hinder the widescale

[33] https://www.loomio.org/
[34] https://opin.me/en/
[35] https://www.airesis.eu/

adoption of these bespoke platforms across Parliament. These will be discussed further in section 8.2.

The rest of this chapter will outline the reasoning behind the use of Discourse and the set-up of the experiments, and the text mining results obtained with the data from both Discourse and Twitter platforms.

The demonstration test was created to compare different modes of engagement across different platforms. The aim is three-fold; (1) to trial a new discussion platform with features aimed at encouraging deliberative discussions; (2) to work alongside parliamentary staff in different areas of parliament and reduce their internal barriers to online engagement such as time-constrains and analysis skills; (3) to observe whether participants would respond differently to the same topic on different platforms.

While a main objective of these experiments was to test a new platform, encouraging data sharing between different parliamentary departments as identified in the digital engagement organigram (Chapter 4) was another key objective. With this in mind, I also worked with the Petitions Committee in order to be able to plug into relevant e-petitions and their supporters. Thanks to this, the Committee agreed to email all signatories of petitions that were in the same topic areas as the inquiries included in my demonstration tests and inform them about the discussion on Discourse. This meant that we were able to reach out directly to interested audiences so they too could have an opportunity to participate and engage further. Following their involvement with the Discourse discussions, several participants went on to submit formal written evidence to the inquiry. Some users also found the new platform through Facebook or Twitter. This shows certain individuals transitioned from an ‘information’ stage of engagement (from following a social media account) to a ‘participatory’ stage (participating in the discussion and submitting formal evidence), as a result of these small experiments demonstrating how parliamentary engagement can manifest in different ways once people are given appropriate methods to do so. Within the second aim, I address the internal barriers to engagement specifically the increased workload experienced by monitoring and analysing thousands of comments received through the online engagement sessions. I hope this will create a greater understanding of how digital tools can be used in parliamentary public engagement, and how the public respond to different features.

6.5.1 Online discussion tools

There are many different online engagement platforms offering a plethora of features depending on the need. There were several options for discussion tools that could be used for the demonstration test, however I decided to use Discourse.

Discourse is an open-source online discussion platform designed to improve the functionality of a traditional discussion forum. It achieves this through its many features such as voting for topics, the option to promote popular topics and polls, login directly through social media, and a “natural immune system” to combat trolls and spammers through moderation and trust systems (Discourse.org, 2019). I set up the Discourse platform as follows: Category > Topic > Posts, where each committee has its own category, and the users organise their posts into topics depending on subject matter. As soon as a user posts a comment, it is automatically uploaded and viewable by other users. Users have the option to make a post as a response to another user’s comment and the original poster can be notified through email or through the app that someone has responded to them. Participants can also create a separate topic within the same category if they would like to discuss something which has not yet been raised by anyone else.

The Discourse platform enables users to moderate themselves by liking or flagging other users' comments and deleting their own comments. If a post is flagged[36], it is sent to a queue which moderators (in this case myself and the committee specialists) can view and decide what measures to take. Discourse also runs a trust system,[37] where new users begin with a trust level of 0, which increases as they read and interact with other users' posts. A higher trust level gives the user access to more advanced features of the platform such as posting attachments, flagging posts, and sending private messages. This is intended to reduce users spamming the platform when they first register, and to give them an opportunity to read and understand the discussion before posting. Users can work their way up to trust level 3, gaining more features as they go, such as the ability to create new topics, unlimited posts and the ability to self-moderate by flagging other users' comments as inappropriate. This is useful for seeing how engaged users are with the site. If a user is being particularly disruptive, I can send them a warning, or remove their ability to post comments or start new topics on the site.

In the past, when the discussion forum on the UK Parliament website was in use, the public were free to post their comments to a specific question, but these needed to be pre-moderated by committee staff before they were published online. This had negative effects for both parties; it greatly hindered the flow of discussion between participants who were unable to respond to each other promptly and sustain a regular conversation; and committee staff were required to read all submissions straight away which used up valuable time and effort (Leston-Bandeira and Thompson, 2017). Therefore, trialling a new platform which had a novel approach to moderation was a key requirement.

Parliament is also keen to make use of existing communities within their online engagement strategies, so having a platform where users could migrate easily from existing online communities such as Facebook was equally important (Liaison Committee, 2015). This feature also worked to reduce the barriers to engagement on the public side in terms of needing to create a new account specifically for the new platform (Nesta, 2019). The set-up and use of the platform was another key factor in choosing Discourse over other platforms. The servers are based in the UK, fulfilling the requirements of Parliament data protection and privacy policies, while also being relatively quick and easy to set-up. This is very important as the business of Parliament is known at most two weeks in advance, so any new software must be equally as flexible and easy to use.

All in all, Discourse was chosen due to its suitability in meeting the demonstration test's requirements in terms of the features it offers, whilst also fulfilling the requirements of Parliament's data security and usability. Any external software used must have its servers based in the UK, and while there are many online platforms, they are mainly based in continental Europe or the Americas where they were originally developed. Having servers in the UK meant that any data was kept within the UK and did not compromise the rules of Parliament data service.

6.5.2 Preparation of demonstration tests with select committees

The demonstration tests were planned and organised in collaboration with select committees in the UK Parliament. The three committees and inquiries participating in the demonstration test were Transport Committee with the Pavement Parking inquiry (Transport Committee, 2019), Environmental Audit Committee (EAC) with the Invasive Species inquiry

---

[36] https://meta.discourse.org/t/so-what-exactly-happens-when-you-flag/275

[37] https://blog.discourse.org/2018/06/understanding-discourse-trust-levels/

(Environmental Audit Committee, 2019), and Environment Food and Rural Activities Committee (EFRA) with the Plastic Food and Drink Packaging inquiry (Environment Food and Rural Affairs Committee, 2019).

The Pavement Parking inquiry concerned the impact of pavement parking on various sectors of society and what measures of enforcement should be taken to prevent ineligible pavement parking (Transport Committee, 2019). The Invasive Species inquiry aimed to "consider the impact and threat to biosecurity from invasive species." (Environmental Audit Committee, 2019). This inquiry also ran alongside Invasive Species Week 2019 from the National Biodiversity Network making it a very topical issue which gained extra popularity as a result and something I wanted to capitalise on during the demonstration test. The Plastic Food and Drink Packaging inquiry sought to understand the current state of food and drink packaging in the UK from both the consumer and manufacturer point of view. The Committee proposed various options for schemes and taxes to see whether this would encourage more of the public to seek alternatives to plastic or increase recycling efforts (Environment Food and Rural Affairs Committee, 2019).

These three inquiries each addressed an issue which people are passionate about or that the public are directly affected by. EAC had also previously been identified as a committee using innovative methods to engage online with the public (Liaison Committee, 2015) and their involvement with the demonstration test supports this. Plastic pollution is also a very popular and important issues, so the inclusion of the EFRA inquiry ensured the demonstration tests had a range of issues that could reach a wide sector of society.

The discussions took place between May and June 2019 and each lasted around 10 days. An overview of the demonstration tests can be found in Table 14. I met with committee specialists in all three committees to review the terms of the demonstration tests. Research suggests digital forums should last for several weeks or months at a time, to encourage deliberative discussion and meaningful solutions (Coleman and Gotze, 2001). However, the committees had other inquiries running at the same time and could not commit to an extended online discussion. As an aim of the experiments was to conduct online engagement sessions without increasing the committee's workload, I was very cautious not to introduce a time-consuming new process. We therefore agreed on the timescales for each demonstration test based on the availability of each committee, how the discussion would be set up on the platform, and how the results would be disseminated and used in the inquiry. This was a crucial step in the process as we were being explicitly clear on exactly what type of engagement we were conducting based on the dimensions examined in 0 and agreed this was to be a consultation exercise. The intention was for the UK Parliament to be the initiators by proposing the discussion, while the citizens would provide their views. This ensured we did not falsely market this activity as a two-way conversation whereby the parliament would be actively participating in the discussion aside from the initial guidance of the terms of reference of the inquiries. This was in important issue to confirm because we could accurately manage the expectations of participants by being clear on how the engagement activity would work. Participants were therefore not expecting the participation of the institution and did not express frustration as a result. For future discussions, a longer period of time for participants to discuss issues important to them would be advised, however this may not be possible in the context of the House of Commons due to the very quick political timings.

Prior to the agreed start-date of each demonstration test, I communicated with the relevant committee to upload between three to four topics on Discourse. Some of these were taken from the inquiry's Terms of Reference and consisted of questions on different issues specific within the inquiry which the committee wanted further views on. The inclusion of these 'seed' topics also gave participants something to start with, when they first logged onto the platform. Seeding the discussion in this way is a good way to begin the conversation and encourage users to participate (Parycek et al., 2014). Participants were able to create their own topics should they feel that there was another issue not covered by the seed topics. As Table 14 shows, participants in the EFRA 'Plastic Packaging' discussion created many more topics compared to Transport and EAC, despite starting with the same number of seed topics from committee staff. The committees were very keen for the public to be able to create their own topics and steer the discussion in a direction they felt was more relevant to them. This power-sharing between the public and the Parliament creates an environment where the institution is open to recognising the views of the public even when they may differ from its own preconceptions.

Alongside meetings with committee specialists, I decided to use e-petitions as a source of promotion for the discussions on Discourse. I identified e-petitions created within six months prior to the experiments that aligned with the same subject matter of the inquiry questions. These e-petitions were chosen manually based on their topic similarity to the inquiry and the petition date. Only petitions in a six-month window could be used because the Petitions Committee have a limit on how long they are able to keep petition signatory data. I shared the link of the Discourse platform with the Select Committee for Petitions, who contacted all signatories of the petitions identified[38] to inform them of the Discourse discussion and inviting them to participate. This ensured that those members of the public who had previously shown interest in a particular topic were not forgotten. They were given the opportunity to participate further, and in a different way to their original act of signing an e-petition. At the same time, the committee Twitter accounts began to tweet the same questions as the seed topics using a specific hashtag. This hashtag was tracked using the Twitter StreamingAPI so any tweet using it could be captured for analysis (Twitter, 2019). It was important to hold the Twitter discussion simultaneously to the Discourse discussions, in order to reduce the number of differing variables in the demonstration test.

One reason was to notify people who have already engaged in some way with Parliament and encourage them to participate again. The second reason was to link data sources within Parliament. Section 4.2 highlighted the different teams with a remit for digital engagement, but also raised the problem of data silos and sometimes a lack of cooperation between departments. The Petitions Committee has a valuable source of data from thousands of UK citizens who have signed an e-petition. This data can be used by other teams for various exercises and I wanted to use these demonstration tests as an example. Sharing data in this way also ensures different areas of Parliament are aware of the discussions which helps teams share best practice engagement methods.

Each committee had two staff members as administrators on the Discourse platform, who could view all comments and moderate any comments flagged by participants. Additionally, I monitored the discussions on Discourse throughout the duration of the demonstration test and when required, dealt with comments which were flagged as inappropriate by other participants. Discourse provides different options for addressing flagged comments depending on the severity of the complaint. If many users flag the same comment, it is immediately hidden from others on the platform. If only one or two participants flag a comment, the administrators of the discussion can decide what action to take between ignoring the comment, hiding it, contacting its author, or suspending them from the platform for a set time period. In the majority of cases (80%), the flagged comments were hidden by administrators (in this case myself and committee specialists). For the most part, these were comments insulting another user or comments completely off-topic. In two cases, the author of

[38] Transport committee petitions: https://petition.parliament.uk/petitions/232684, https://petition.parliament.uk/petitions/222715, EAC petition: https://petition.parliament.uk/petitions/235425, EFRA petitions: https://petition.parliament.uk/petitions/232684, https://petition.parliament.uk/petitions/222715

the post was suspended from the platform for making repetitive posts or creating a new account using the same IP and email address. This feature allowed administrators to avoid trolls as much as possible on the platform, while responding to the concerns of the existing participants. This feature was one of the primary reasons for choosing Discourse and became very useful during the EAC discussion which received 75 (11% of total comments) flagged comments (Table 14). This is not a feature available on Facebook or Twitter and can help to ensure the discussion remains on track and remains appropriate.

As mentioned earlier, Discourse also features a trust system, where a user starts with a level of 0 and can then move up as a reward for good engagement obtaining additional abilities to shape and moderate the debate. By the end of the experiments, 38% of users on the platform were trust level 1 or higher and had spent an average of 44 minutes on the platform reading different comments. This gives clear evidence that users were using the platform well and spent time to read and respond to the opinions of other participants. The committees were able to use this to clarify the submissions being made to their inquiry were being discussed properly and not just a result of users posting without reading comments from any other users.

On the previously agreed end-date of the discussion, a committee specialist made a post on the platform thanking all the participants for their comments, and explaining how their submissions would be used in the inquiry. They also provided links which participants could use to follow the rest of the inquiry. This action of closing the feedback loop with the participants and acknowledging their time and contributions was an important aspect of the demonstration tests. Several participants also expressed gratitude for the opportunity to contribute, for example “Thanks for the opportunity to comment” and “Thank you. I will follow the development of the inquiry with interest.”.

At the end of each demonstration test, I created an interactive html report for each committee with detailed information about the primary topics, sentiment analysis results, locations of participants, and comparisons between Discourse and the Twitter discussions. Committees also received a csv file of all comments posted to the platform, so they had a record of all comments in raw format.

Table 14: Summary of demonstration tests

	Committee		
	Transport	**EAC**	**EFRA**
Inquiry	Pavement Parking	Invasive Species	Plastic Food and Drink Packaging
Discussion Dates (2019)	10 - 17 May	14 - 20 May	6 - 13 June
Number of Discourse Topics	3	7	142
Number of Comments	108	665	2946
Number of Participants	95	322	980
Comment Flags	0	75 (11%)	2 (<1%)
Twitter Hashtag	#PavementParking	#EACInvasiveSpecies	#FoodPlastics

Number of Tweets	461	82	284
Number of Twitter Users	302	52	266

6.6 An assessment of platforms and discussions

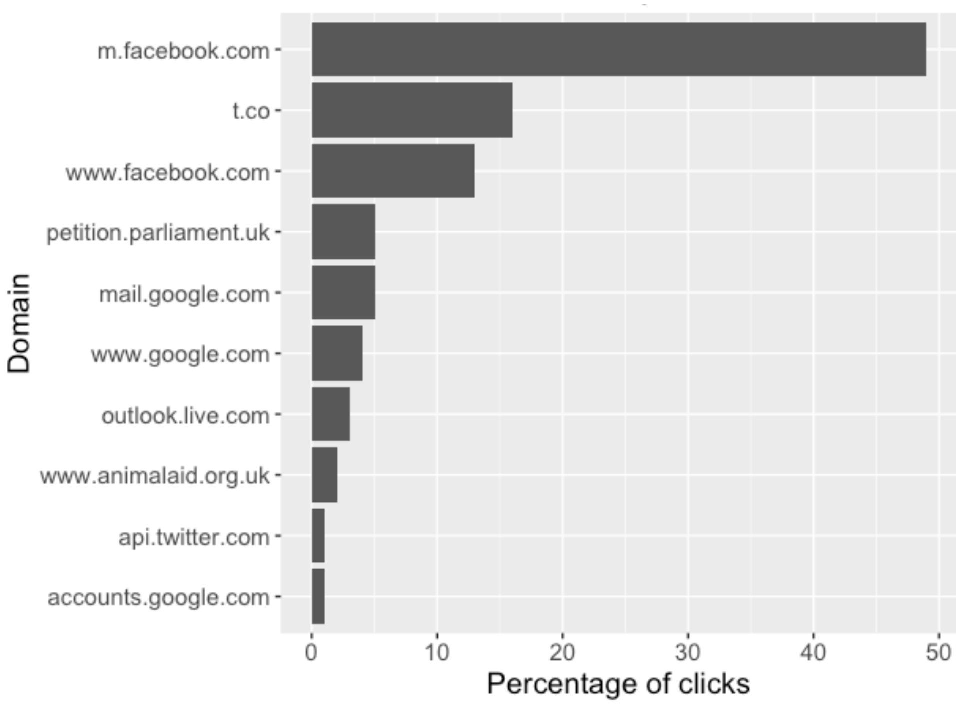
A main objective of these experiments was to explore whether using a new platform with features designed specifically to enhance online discussions would make a difference to the way participant interact with the UK Parliament. The main goal of the demonstration tests was to increase the quality of citizen input, and to facilitate the inclusion of this citizen input into policy making in Parliament. One way to influence the quality of the debate is through trying a new platform which was created specifically for use in online discussions and online communities. This section will provide results from the computational analysis of comments posted in the demonstration tests as well as social network analysis results exploring the interaction between participants during these demonstration tests. The analysis of comments and social interactions on Discourse will then be compared with the comments and social interaction on Twitter that serves as a quasi-control group in these demonstration tests. This allows me to understand the benefits of various features that Discourse offers and that are expected to enhance the effectiveness of online inquiries and debates initiated by the UK Parliament.

6.6.1 Discourse Platform and Twitter Activity Analysis

6.6.1.1 Where did the participants come from?

Understanding from which website participants transition to the site of online engagement helps the teams within the UK Parliament coordinate their future engagement activities. They have a clearer idea of the sources and channels participants use to find their engagement activities, as well as insights into which organisations or charities are also directing members of their communities to online engagement sessions. The majority of the traffic to the Discourse site for all three discussions came from Facebook (m.Facebook.com refers to the mobile version of the site suggesting many users were using their mobile phones to participate rather than their computers) and Twitter (t.co), followed by the UK Parliament e-petitions site as shown in Figure 57. This pattern persisted for the Transport and EFRA discussions, however for EAC Twitter was the most popular linked site. 6% of users to the EAC discussion came from the not-for-profit company AnimalAid.org.uk who aim “to work, by all peaceful means, for an end to animal cruelty”. This site also had a link to an e-petition to make squirrels exempt from the Invasive Species act and to the Discourse forum with advice on what points to raise in the discussion (AnimalAid, 2019).

Figure 57: Discourse traffic sources – all demonstration tests



In Chapter 5, I explained that one difficulty of using social media as a data source was the lack of quality information regarding location data, specifically in tweets. While in Twitter data we rely entirely on the user to provide that information in their profile description, location information is not at all available on Facebook. However, geo-spatial information could be valuable to understand where in the country participants are located and whether a topic is of interest across a country or has a regional focus. Discourse differs from Twitter and Facebook as it allows me to infer reliable location information from users' IP addresses, from which latitude and longitude can be derived and plotted, so I can explore where in the country participants are based (see section 3.3 for details). This is a useful form of information for select committees, whose general goal is to engage with people from different areas of the country, but who might also be interested to understand in which parts of the country citizens are particularly concerned about a certain issue. Having an accurate location tag enriches the analysis of online discussions and, depending on the granularity, can pinpoint which constituency members are impacted by which issues.

The majority of participants across the three Discourse discussions were located in the UK, however while EFRA (Figure 60) and EAC (Figure 59) discussions favoured London as the city where most participants were based, and most participants in the Transport (Figure 58) discussion came from Plymouth. It is not surprising that London is a hotspot of engagement, it is large city with a very diverse population, many citizen interest groups, non-profit organisations, associations etc. are located there and it is the centre of British politics. Still, other major cities such as Brighton, Manchester, Nottingham, Bristol, and Leeds had participants across the three discussions. In terms of the spread across the UK, both EAC and EFRA discussions saw slightly more participants from the south of England than anywhere else in the country, while Transport had a more uniform distribution (albeit with far fewer users than the other discussions). In all cases, Scotland, Wales and Northern Ireland had far fewer participants than England. This may be because Scotland, Wales and Northern Ireland all have their own devolved parliaments who conduct their own online engagement sessions focussed on their populations. The UK House of Commons aims to represent all British citizens, not just the English, but this finding shows that they are not reaching those in the devolved regions very well.

Figure 58: Transport Committee 'Pavement Parking' Discourse geographic distribution of participants. Map points sized by number of participants.

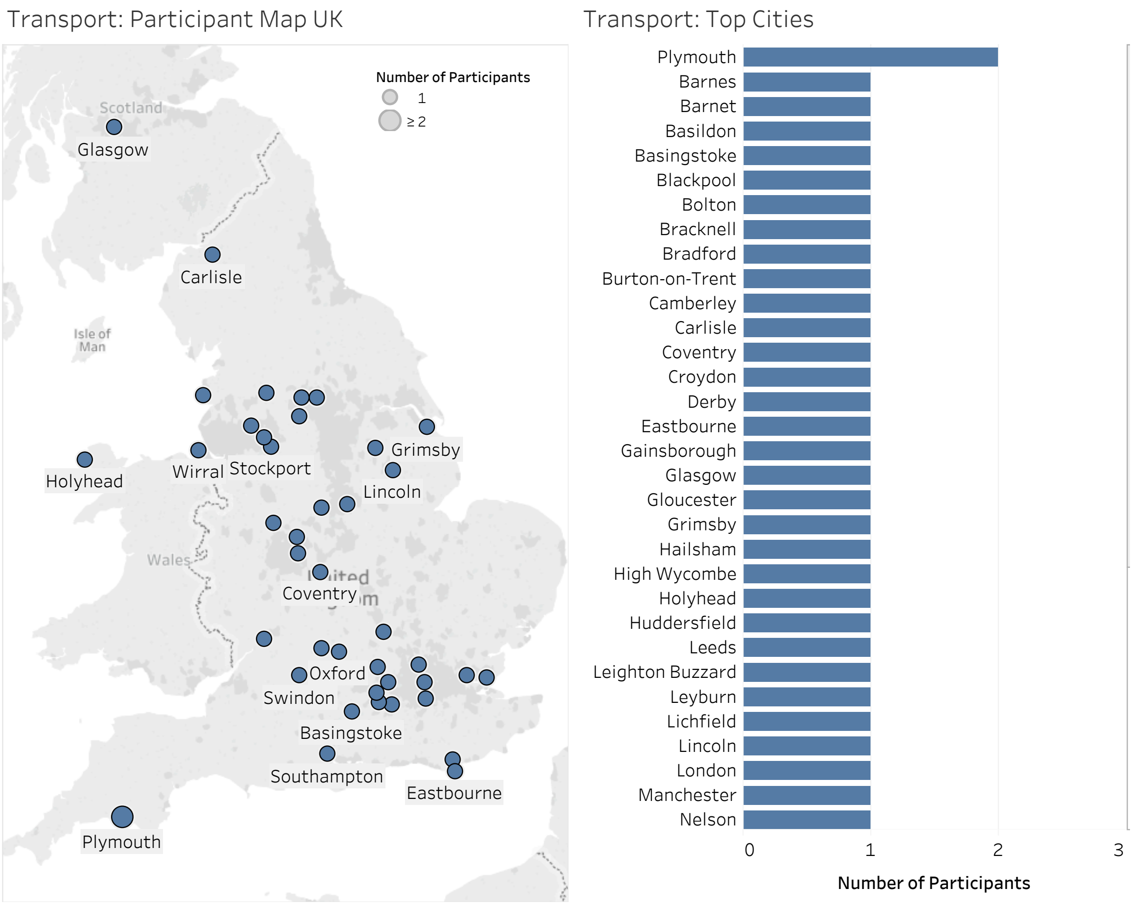
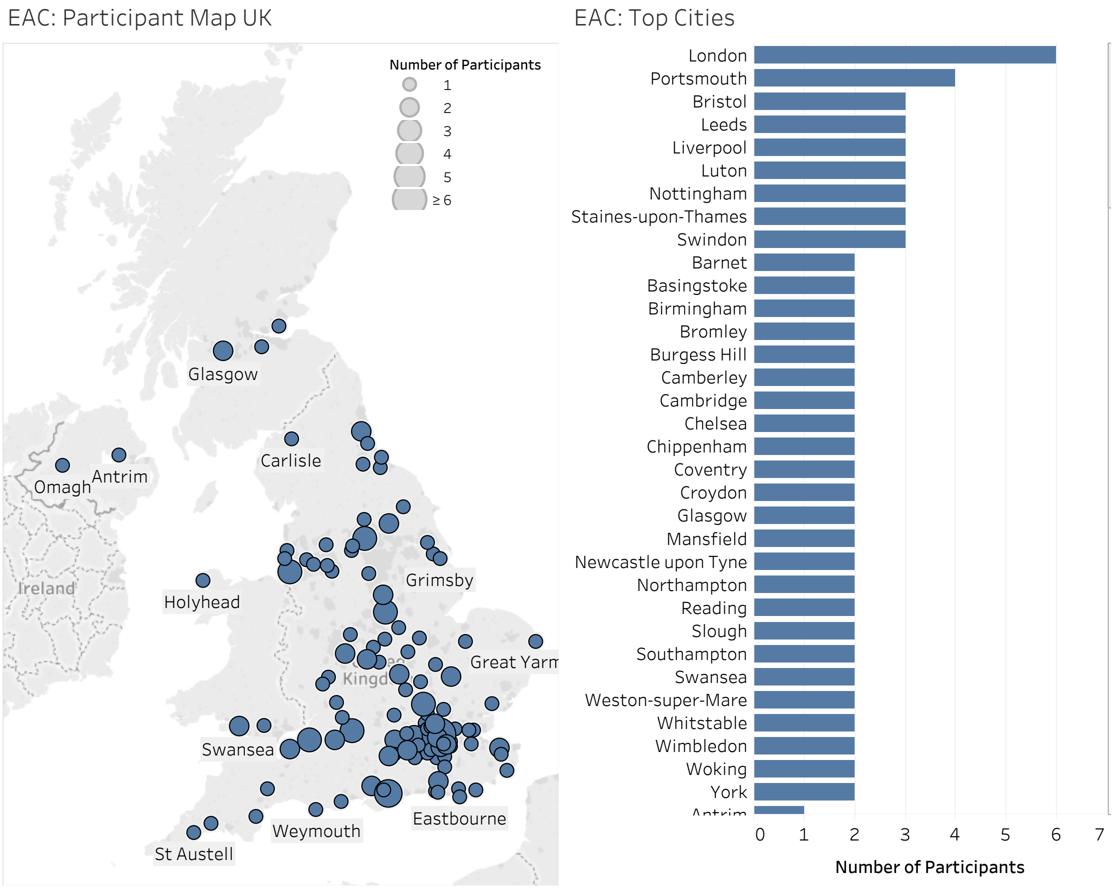


Figure 59: EAC Committee 'Invasive Species' Discourse geographic distribution of participants. Map points sized by number of participants.

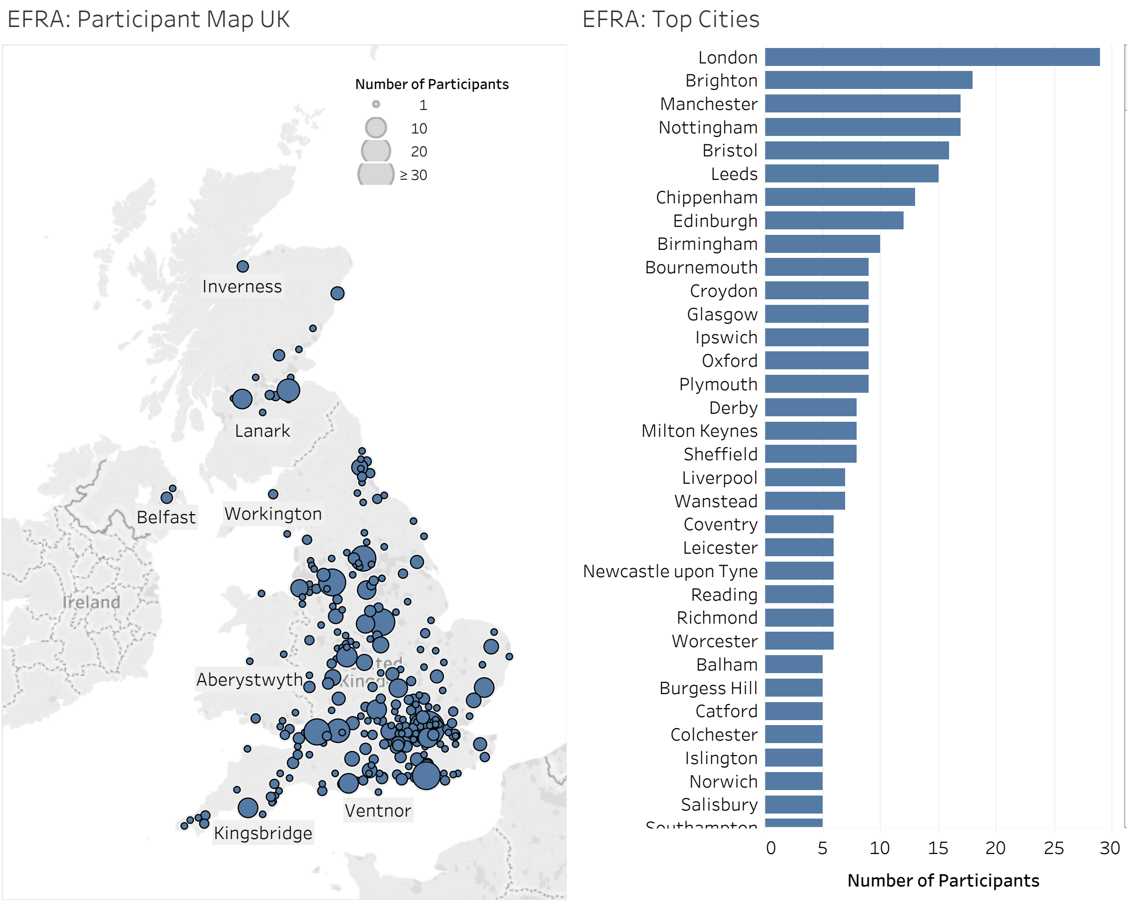


This demonstration test did not target any particular geographical region, instead aiming for a more organic audience. However, these results give insight into the reality of online engagement and the areas of the country that are affected and interested in these topics. As the Discourse users were partly derived from petition signatories, there could be a link between the geographical spread of Discourse participants and signatories of the related petitions (Unboxed, 2019). I have investigated the link manually by examining the number of e-petition signatures for different locations in the UK and comparing them to the Discourse location results. While there was some overlap between Discourse participants and e-petition signatories in the EFRA discussion in Falmouth, likewise with Transport users in Plymouth, and EAC participants in the south-east of England, these are not enough to firmly suggest that the geographical spread of Discourse users was greatly influenced by the e-petitions. Therefore, I must explore other reasons for the pattern of user locations.

The Transport Committee's pavement parking inquiry saw the most uniform distribution of participants across the country, suggesting this issue of cars restricting the use of the pavement and public footpaths is one that is felt in many areas in the UK. However, as previously mentioned, the low numbers of participants in this discussion makes it difficult to draw any concrete conclusions. Conversely, the discussion on invasive species, specifically on red and grey squirrels showed clusters of users in the south especially Greater London, Portsmouth, and Bristol. Whereas grey squirrels can be found all over England since their introduction in the 19th century (Welsh Government, 2018), research suggests the majority of red squirrels that are existentially threatened by the presence of grey squirrels are found in Scotland, but also Dorset and Northumbria in England (Countryfile, 2019; Wildlifetrusts.org, 2019). These areas however are mostly missing in the location map for the EAC discussion, that only includes a small cluster of users in the North-East of England. This suggests that people in the regions mostly affected by the grey squirrels as an invasive species have not really come forward to participate in the debate. On the other hand, there are many other invasive species that raise concerns across the UK, e.g. Japanese knotweed, Asian hornet, American bullfrog etc. Therefore, in terms of squirrels, the committee was not necessary hearing from those who were directly affected by this issue (just those who were perhaps most passionate), but given the discussion topic was set out to be broader including all types of invasive species, the opinions of citizens from across UK were of interest.

Unfortunately, I could not compare the geographical distribution between the two platforms used in the demonstration test because of the unreliable location information on Twitter provided by users. Nevertheless, through analysing the available location data from Discourse, the Digital Engagement team can evaluate if they are reaching citizens in different parts of the country and which issues are important to different locations.

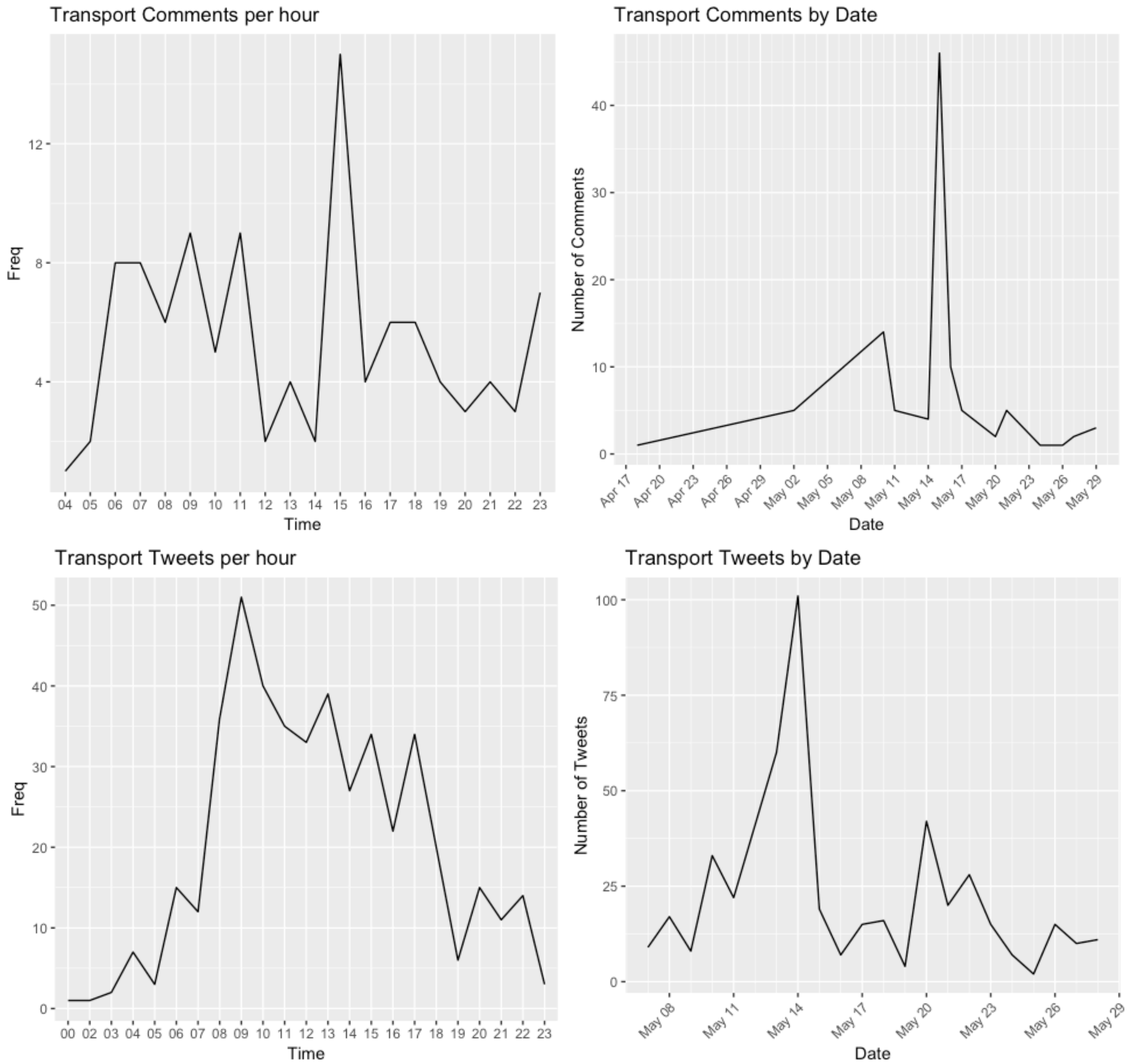
Figure 60: EFRA Committee 'Plastic Packaging' Discourse geographic distribution of participants. Map points sized by number of participants.



6.6.1.2 When did they participate and for how long?

In section 6.2.2 , analysis of the activity of users on Facebook gave insights on their daily habits and provided a richer understanding of the participants. In a similar way, there was also a difference in the timings of posts made between the three discussions. The Discourse discussions for Transport (Figure 61) and EFRA (Figure 63) had a clear peak time of the day when participants were online and most active, 15:00 and 08:00 respectively. These discussions also had a definite peak in terms of the day when most comments were posted, usually happening over a single day. EAC (Figure 62) differed from this pattern in both timings and dates of posts on Discourse. In this case, there was no single spike of comments during the day, but smaller peaks at 10:00, 14:00, 17:00, and 21:00. Participants were also active over several days with two major peaks, rather than one. This suggests a pattern of repeated visits by users several times a day, in which they would respond to a reply on their comment, and returning day after day to continue the conversation. This discussion was also the most controversial and

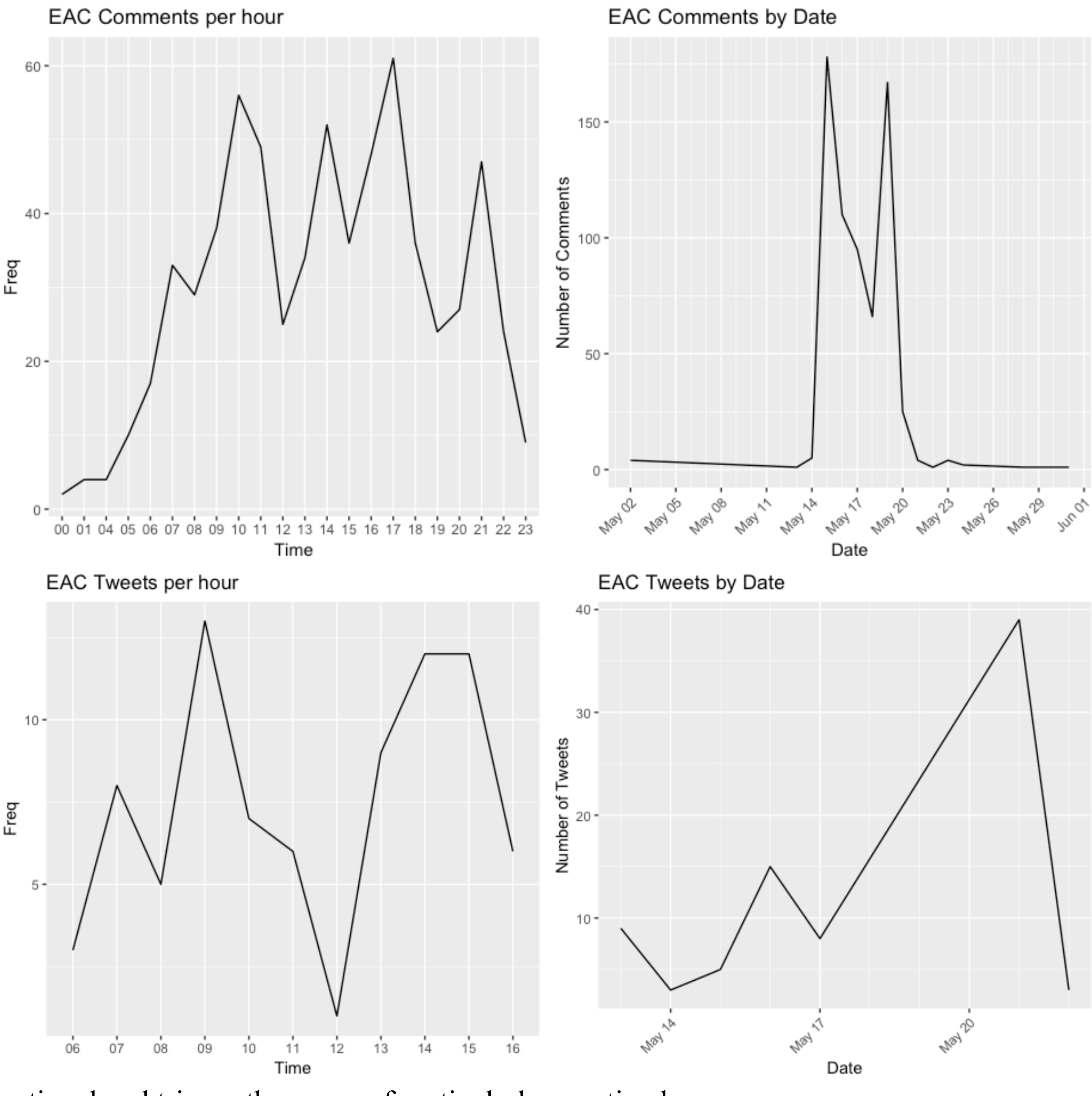
Figure 61: Pavement Parking (Transport Committee): temporal frequency of posts for Discourse (above) and Twitter (below)



negatively framed of the three due to a discussion about red and grey squirrels as will be discussed in the next section.

This pattern of repeated visits to the platform during the day and peaks over several days was also observed in the controversial Animal Fur digital discussion held on Facebook which also featured very negative comments. The link between engagement levels and the negativity (in particular anger) associated with a topic has in fact been established in earlier research (see for instance Weber (2013); Wollebæk *et al.* (2019) and Haro-de-Rosario, Sáez-Martín and del Carmen Caba-Pérez (2018)). My results seem to confirm these earlier findings and could suggest a link between the frequency and pattern of user activity and the sentiments of comments. Interestingly, both discussions (Animal Fur and EAC's Invasive Species) were on the topic of animals. It appears that topics resolving around animal protection are very

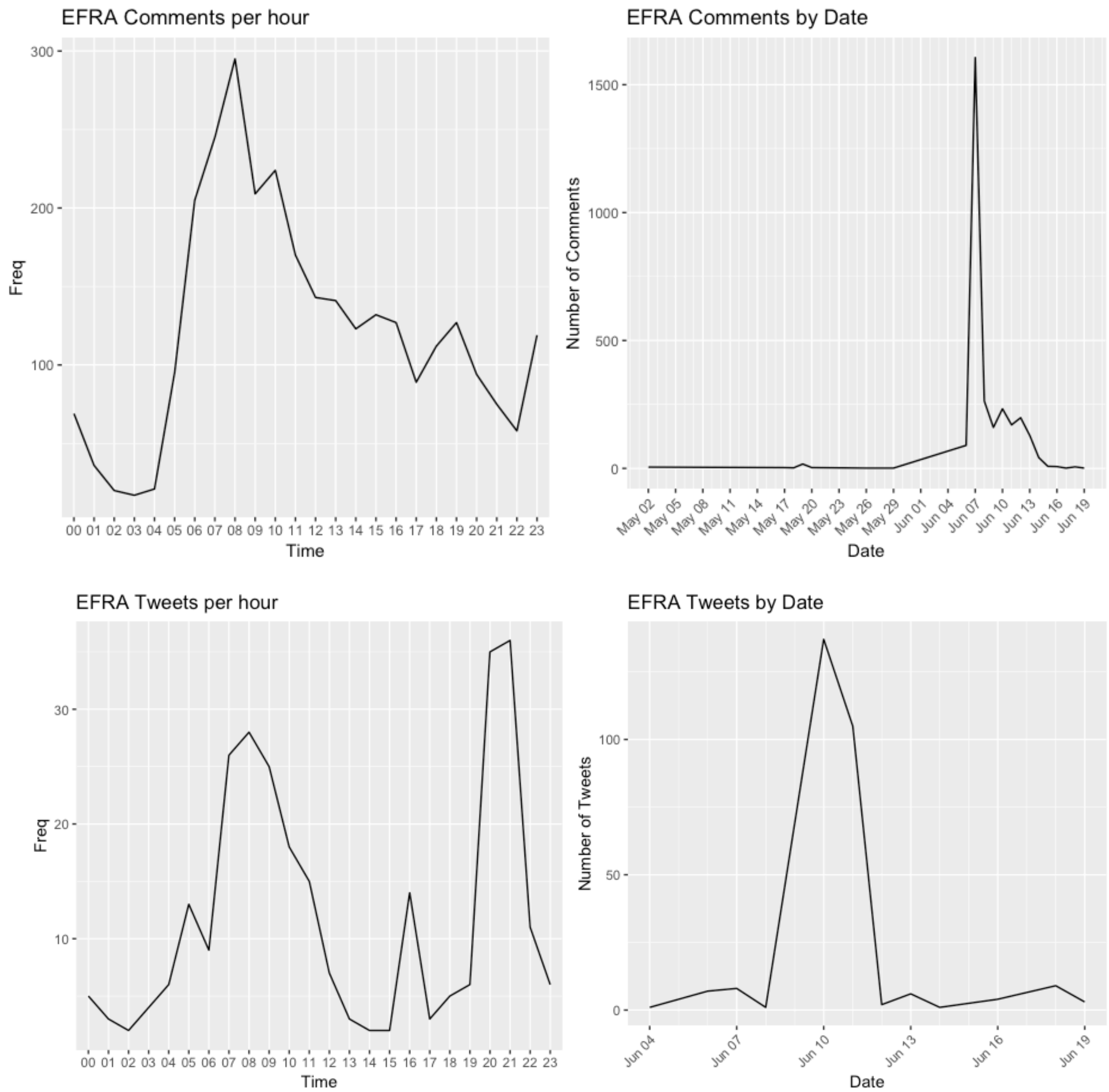
Figure 62: Invasive Species (EAC Committee): temporal frequency of posts for Discourse (above) and Twitter (below)



emotional and trigger the usage of particularly negative language.

On Twitter, EAC had the shortest range of times for users to tweet, generally in the morning and again in the afternoon while the 21st May was the most active day for tweets. This falls after the spike in activity on Discourse and suggests users perhaps continued the discussion on Twitter following the end of the Discourse discussion. The Transport discussion on Twitter featured a slightly earlier peak of activity than on Facebook with most users active around 09:00 compared to 15:00 on Facebook. EFRA's Twitter discussion differed heavily from the Discourse discussion regarding the times users were most active. This pattern has two peaks of activity in the morning and at night, with a small peak during the afternoon. This might suggest that different groups of citizens with different time schedules engaged on the respective platforms. This probably was also the case with the other debates, where I also see differences in the contribution timings between the two platforms. The dates on Twitter however were very similar across the two platforms in the EFRA discussion with one day (10th June) having the greatest number of tweets and very few on other days. As with EAC, this peak day of the EFRA Twitter activity came after the main peak of Discourse activity and suggests the discussions on Twitter were slightly delayed compared to the Discourse forum, or could suggest that users from the Discourse forum continued the discussion on Twitter. This could be because (as with all discussions) the committee staff made a post thanking the participants

Figure 63: Plastic Packaging (EFRA Committee): temporal frequency of posts for Discourse (above) and Twitter (below)



of the Discourse discussion at the end of the demonstration test, letting them know that they will no longer be monitoring the discussion but participants were still welcome to continue the conversation. Many participants could have seen this and decided to stop posting. However, as Twitter is a social media platform with many users posting about similar topic to the inquiry, Twitter users may have felt encouraged to continue the conversation. Furthermore, where users would have received a notification when the committee staff made their post on Discourse, the same is not true for Twitter, so Twitter users may not have been aware the committee was no longer monitoring tweets. Hence one important difference between the two platforms, is that the discussion on social media like Twitter are necessarily much less coordinated and fluid.

6.6.1.3 How did they interact with each other?

The pattern of interaction networks differs heavily across both platforms, with the Twitter networks displaying a *broadcast network* (Smith *et al.*, 2014) pattern comprised of hubs of main accounts surrounded by users who are retweeting that account across all three debates. There is little interaction between the leaf nodes in these types of networks. On the other hand, the Discourse networks across all three debates are much more connected with higher average degrees and network diameters, (see summary of network metrics in Table 15). All networks across both platforms feature low density levels suggesting there could be many more connections between nodes than observed, however the Discourse networks have a slightly higher density than those of Twitter.

Table 15: Discourse and Twitter demonstration tests network metrics

Metrics / Committee	Discourse replies			Twitter retweets		
	Modularity	Weighted Degree	Diameter	Modularity	Weighted Degree	Diameter
Transport	0.68	1.90	3	0.83	1.0	4
EAC	0.42	6.02	13	0.39	1.4	2
EFRA	0.75	3.19	11	0.56	1.0	2

Using the network of replies, the majority of the EAC clusters are well connected with a modularity of 0.42 and a weighted degree of 6.02 (Figure 64). Users with the largest node size have the highest degrees which means they have the most connections, or more specifically, have made the most responses in the discussion. These users user2 and user12, who are the hubs of the two main clusters in orange and green, were primarily involved with the sub-debate on red and grey squirrels raised in several topics (explained in section 6.6.2) and have many connections between each other. Other users outside of the orange and green clusters featured fewer connections and therefore replied to each other on a smaller scale. This can be a result of the different types of discussions in these subtopics. Users discussing other invasive species such as plants and bees did not have as many replies between each other, as users engaged in the debate about the welfare of squirrels. Contrastingly, the same discussion on Twitter (Figure 65) was primarily impacted by the EAC Twitter account (@commonsEAC) which was retweeted many times by various users including the @InvasiveSpecies account. The Twitter discussion also showed many retweets between different users including "wcl_news" (Wildlife and Countryside Link), "ukladybirds" (UK Ladybird survey run by

Helen Roy), and "tthecccuk". However, this network had a smaller weighted degree of 1.4 suggesting users had few retweets between each other and mainly focussed on one central user during the discussion. Most likely, this was users responding to the question from the main EAC Twitter account giving their views on the subject of invasive species. These findings from the network data combined with the previous analysis of activity time patterns show the difference between the two debates on the two platforms. The repeated peak and trough pattern throughout the day activity from the EAC Discourse discussion appears to be linked with a tightly connected network of users who often interact and respond to each other often. Meanwhile, the less-active users on Twitter who just interact at certain times of the day appear to do so only to respond to questions posted by the committee and less so to interact and respond to each other. The Twitter network can be likened to a broadcast network with many users re-tweeting a central account, while the Discourse network more closely resembles a tight crowd network where there are many interactions between different users (Smith *et al.*, 2014).

Figure 64: Network of replies in the EAC (Invasive Species) Discourse discussion. Nodes weighted by degree centrality and coloured by Louvain method partition. Edges weighted by number of replies between nodes

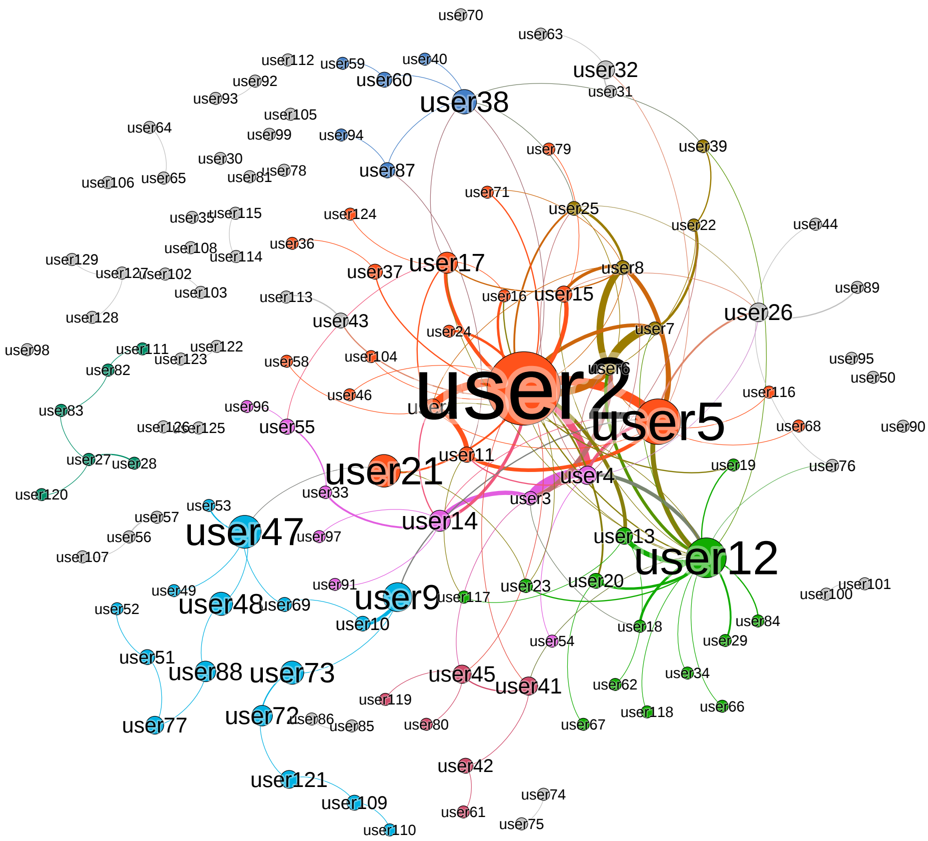
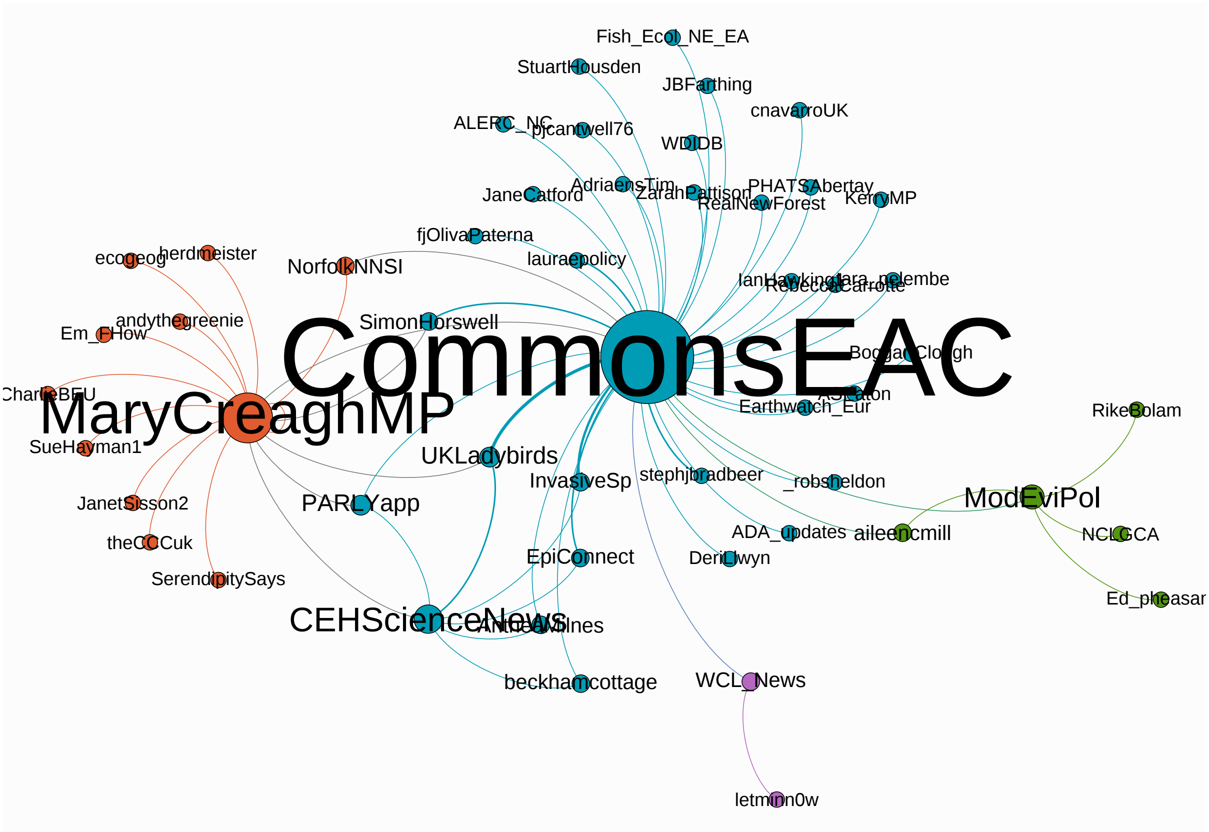
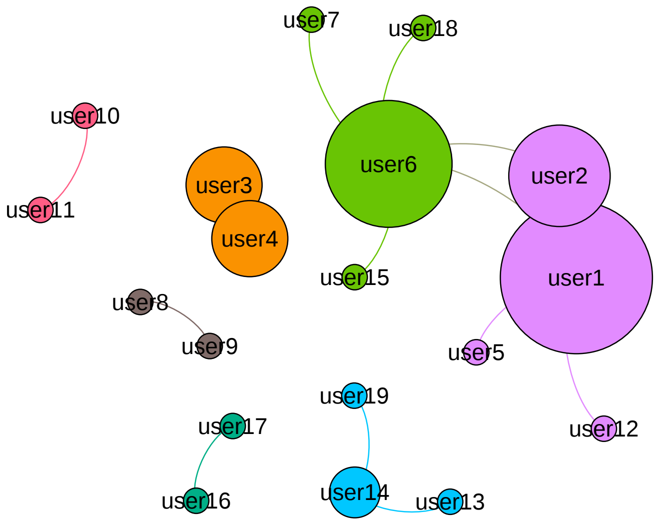


Figure 65: Network of retweets in the #EACInvasiveSpecies Twitter discussion. Nodes weighted by degree centrality and coloured by Louvain method partition



Where the EAC Discourse graph shows many users responding to each other, the Transport Discourse discussion is much smaller and sparse with few edges connecting the user nodes (Figure 66). This discussion had fewer participants, so the graph is understandably smaller with a network diameter of 3 and an average weighted degree of 1.9. The large nodes represent users with high degrees that have responded to the most users, for example user1 has a weighted degree of 6 so has responded to 6 people. This network has a relatively high Louvain modularity of 0.68 which suggests a relatively good separation of nodes into communities, but finds 7 small communities with between 2 and 4 user nodes in each and has a pattern of a brand clusters or tight crowd network (Smith et al., 2014) with small clusters of users with little connections between them.

Figure 66: Transport Committee (Pavement Parking) Discourse reply network. Nodes weighted by degree centrality and coloured by Louvain method partition



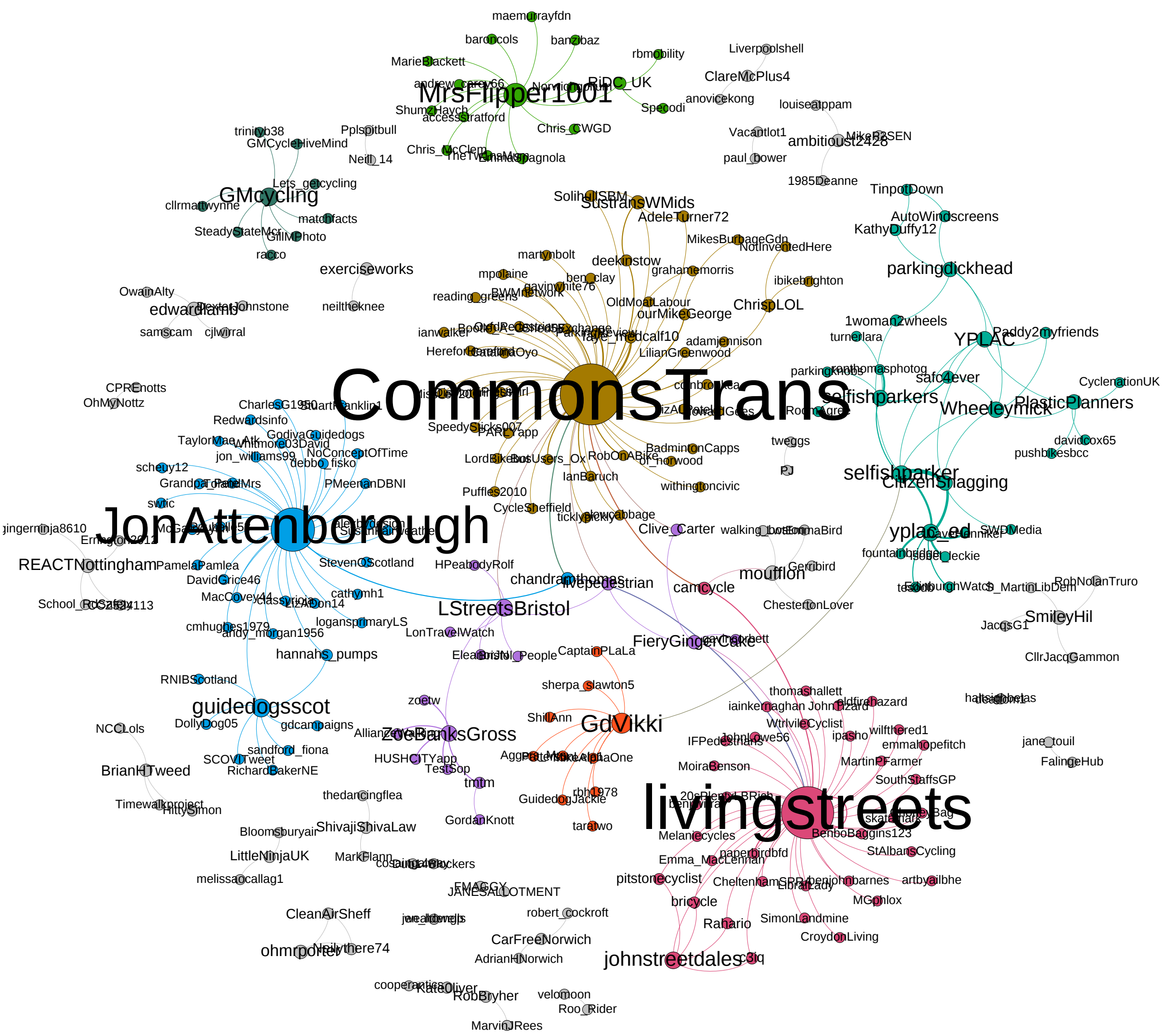
Although Transport had the fewest Discourse comments and participants of the three demonstration tests, the same #PavementParking discussion on Twitter had the most tweets and Twitter users participating. Therefore, where I see differences in the activity and level of engagement across the two platforms in terms of the number of users involved, there is also a difference in their interactions in terms of how those users engage with each other.

The network of retweets for the #PavementParking discussion is shown in Figure 67. This network has a modularity of 0.83 with the largest communities shown in colour. The figure reveals a much more detailed network with three large hubs centred around the @CommonsTrans (Transport Committee), @JonAttenborough (Technology expert and innovator/Disability and inclusion advocate)[39], and @livingstreets (UK charity for everyday walking)[40] accounts. A network such as this can be described as a broadcast network (Smith et al., 2014). Each of these hubs have a sub-community of Twitter users who have retweeted the central account, however the nature of this network means the average degree is only 1.0. This means there is only an average of one connection (in this case retweet) per user node. Comparing this with the Discourse network of replies, although of much smaller graph size, it had an average degree of 1.9, suggesting a slightly more connected network. The network diameter measures the longest of all the shortest paths between nodes in a network. This measure is at 3 for Discourse and 4 for Twitter, suggesting the latter had more paths to travel to reach one end of the network than the other. This would imply a more connected network, however, the high modularity and pattern of nodes surrounding a central hub of the Twitter discussion means that there are fewer edges for these nodes to travel to reach the farthest node. This can also be explained through the broadcast network vs. tight crowd/brand clusters pattern of the two networks. In practical terms, this means that users were more connected to each

[39] https://twitter.com/jonattenborough?lang=en
[40] https://twitter.com/livingstreets

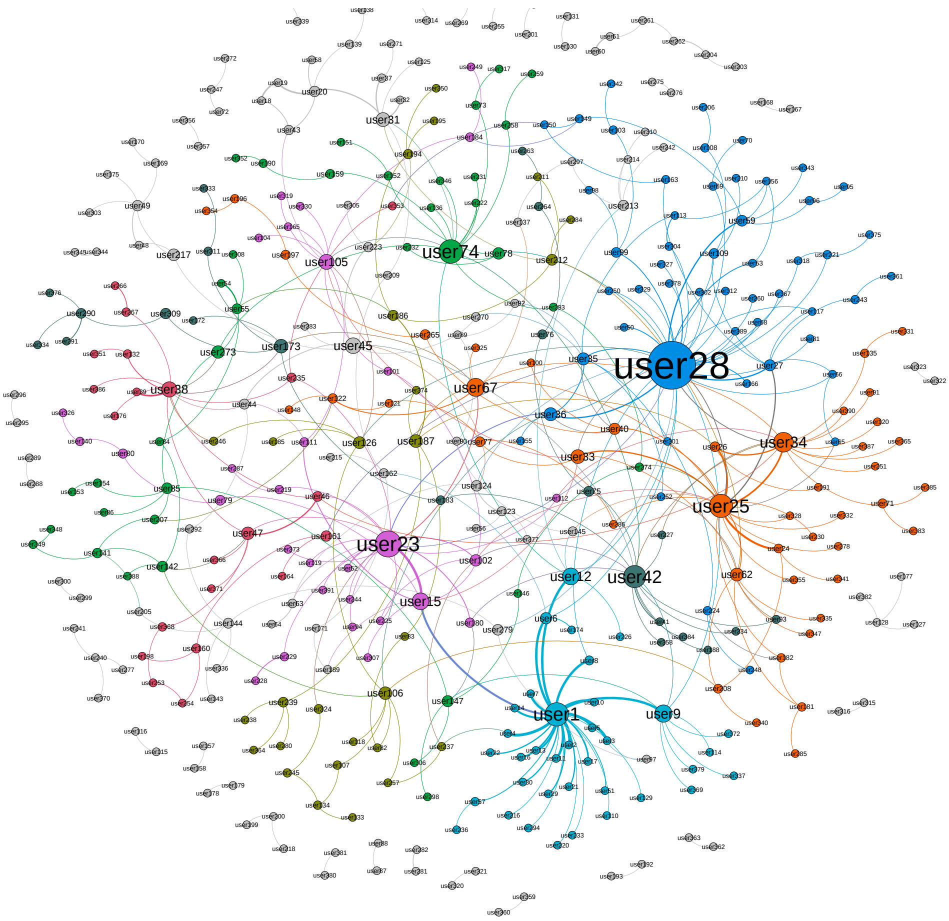
other in the Twitter discussion, but the discussion would be more centred around one main account in each community rather than involve a mix of different users.

Figure 67: Network of retweets in the Transport #PavementParking Twitter discussion. Nodes weighted by degree centrality and coloured by Louvain method partition



The final discussion by EFRA committee on Plastic Packaging also exhibits the broadcast network pattern on Twitter of central hubs surrounded by leaf nodes (Figure 69). @commonsEFRA (official EFRA committee Twitter account) and @Botanygeek (BBC Science presenter and New Scientist columnist) have their own communities in the network, and were retweeted by a range of users who follow them or the #FoodPlastics hashtag which was used in the Twitter discussion. As with the other Twitter social networks in this section, there are few links between these leaf user nodes showing they have interacted only with one main account and less so with each other. There is a modularity of 0.56 however, as these hubs are very self-contained with little connections between them, the network has a small diameter and average degree. However, there are three user nodes in the Twitter network (laurensmckee, ritasymon, and PaivooPiez) who act as cutpoint nodes or bridges between two communities whose central nodes are CommonsEFRA and Botanygeek. These users are the only connection between the two main clusters in the network and unlike other users, have interacted with both hub-accounts. They see content from different groups of users and may have a more well-rounded understanding of the whole discussion. In this case, section 6.6.2 shows that subjects covered in the Twitter discussion included how users could submit further evidence to the inquiry as well as more general issues surrounding food waste and the environment. However, more polarising discussions between communities on social media who do not interact with each other could lead to users themselves becoming more extreme in their opinions. For

Figure 68: EFRA Committee (Plastic Packaging) Discourse network of replies. Nodes weighted by degree centrality and coloured by Louvain method partition. Edges weighted by number of replies between nodes

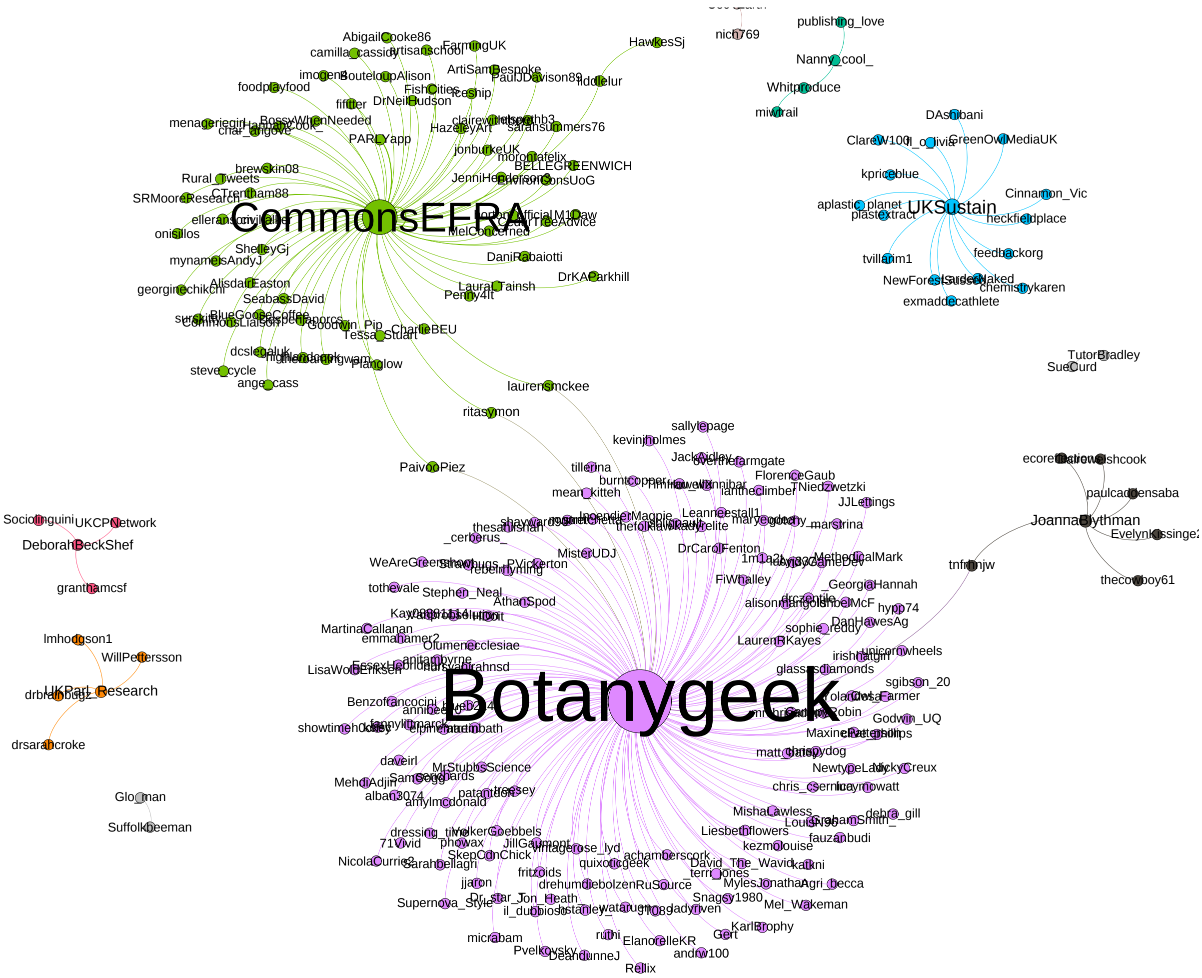


example, had this discussion been about something very divisive such as Brexit and the two main communities being for and against respectively, the three cutpoint nodes would be the only users hearing both sides of the debate.

The same discussion on Discourse features a high modularity of 0.75 and a very connected network with shared edges between various user nodes (Figure 68). The network has a weighted degree of 3.19 meaning each user node has responded to around 3 other users on average compared to just 1 on Twitter. In this social network, there are less defined communities of users, because there are many connections between them suggesting all users in this discussion communicated well with each other within the different sub-topics. This was the largest discussion in terms of topics and users on Discourse (Table 14), however the network shows that this did not cause the users to splinter into different communities with little communication between them.

While this analysis shows similarities among twitter networks being largely unconnected with each other with users instead focussing on one single Twitter account. Insights such as this give the Digital Engagement team some understanding of how their engagement activities are received in different environments. As mentioned earlier, the primary aim of these demonstration tests is to increase the quality of citizen input. This can be defined as input which enables citizens to communicate effectively with each other and provide well-discussed opinions to Parliament. Keeping this in mind, the social network analysis in this section suggest Discourse is a more appropriate platform for participants to discuss issues important to them and crucially interact with each other.

Figure 69: EFRA Committee (Plastic Packaging) twitter network of retweets. Nodes weighted by degree centrality and coloured by Louvain method partition.



6.6.2 What did they have to say and how did they express themselves?

So far, I have explored the users' activity on the two platforms of Discourse and Twitter across the three committee discussions. This gives an insight into where the participants came from and how they interacted with each other but does not provide information about what they were saying or how they said it. For this, I need natural language processing to study the differences between discussions. In the context of the experiments, I have two dimensions of comparison; between platforms and between committees. While the inter-platform comparisons are the primary focus in this section, I also explore the differences in discussions and how the varied subject matter of the experiments contributes to the patterns I find in the analysis. Exploring differences between platforms allows me to assess how people react to the same topic and how (or if) the design of the platform can influence these discussions and reactions. In the context of parliamentary engagement, an understanding of these differences is crucial to gain the most out of online discussions. Previous research in this field has found that users act differently depending on the channel of engagement they are using. For example, the MN-Politics forum was developed in Minnesota, USA and had two types of forums; real-time chatrooms and asynchronous email lists and bulletin boards. Users of the forum engaged in more small talk and were more laid back in their responses in the real-time chatrooms, but were more serious and engaged in rational debate in the email lists (Smith, 2009; Jensen, 2003).

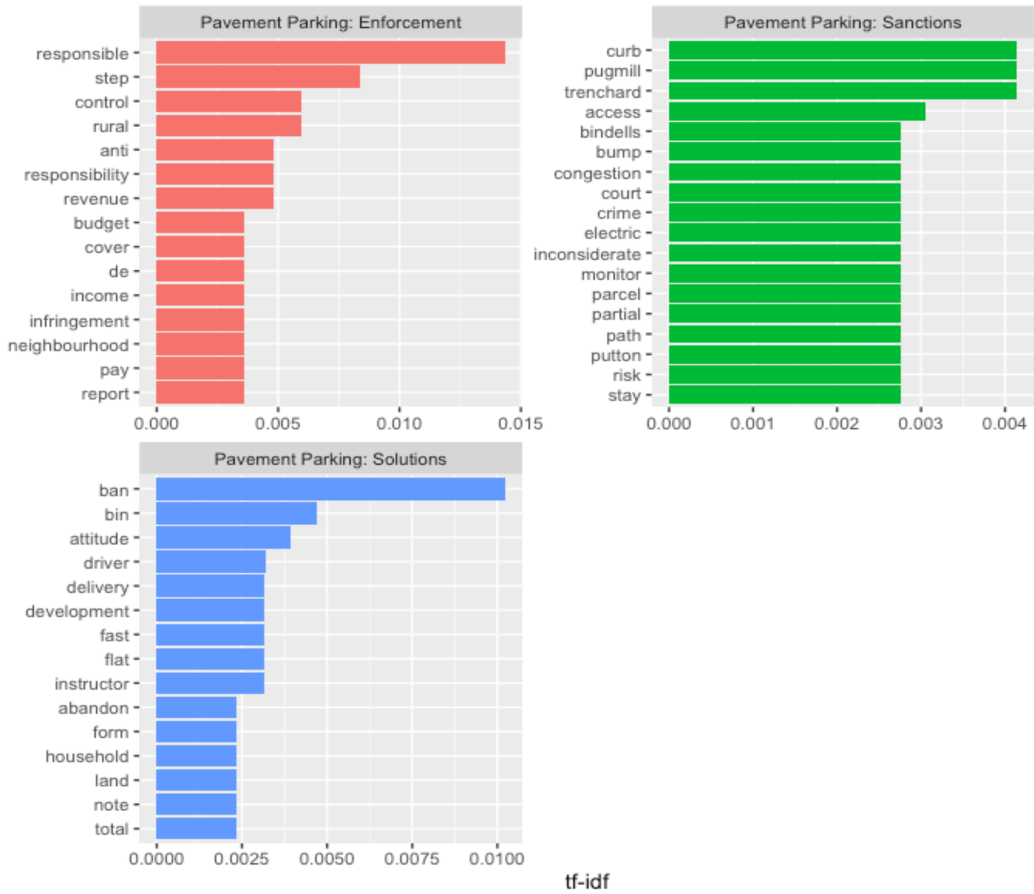
Table 14 shows there was a great variation in the numbers of user-created topics across the three discussion, with EFRA inquiry on plastic food and drink packaging having the most by far. For the EFRA discussion, LDA topic modelling was used to group the 142 user-created Discourse topics based on the similarity of words used in each (Blei, Ng and Jordan, 2003). The optimum number of topics in each LDA model was decided using the LDA Tuning package in R (Nikita, 2016). In the case of EFRA's 142 topics, the model decided on 12 being the optimal value of *k*. I then used the posterior probabilities of this model to categorise the EFRA Twitter discussion into the same topics. The *tf-idf* measure was then used on the corpus of words in each topic to identify the words most specific to each topic to explore the main issues raised in these discussions. As the Transport and EAC Discourse discussions had a manageable number of user-created topics, no topic modelling was used and I instead used text mining analysis to examine which words appeared in each the user-generated topics. However, for the respective discussions on Twitter, I also used LDA topic modelling algorithm to determine the number of topics and examine them in relation to the user-defined Discourse topics.

The National Research Council of Canada (NRC) emotion lexicon was used to identify the sentiments of the discussions across both platforms and across the three discussions. This lexicon was chosen in agreement with the select committees who indicated the different categorisations of emotions along with the positive and negative distinction included in the NRC lexicon was a feature they were keen to explore in these demonstration tests. Throughout the three demonstration tests, the Discourse discussions had a much more varied sentiment category distribution than the discussions on Twitter. Twitter discussions had more polarised sentiments with the majority of comments expressing between three to four main sentiments. In all discussions on both platforms *positive* sentiments were the majority of sentiments expressed, however there are some differences in the discussions and platforms regarding the other sentiments, which will be discussed in more detail in the following sections.

6.6.2.1 Transport Committee 'Pavement Parking'

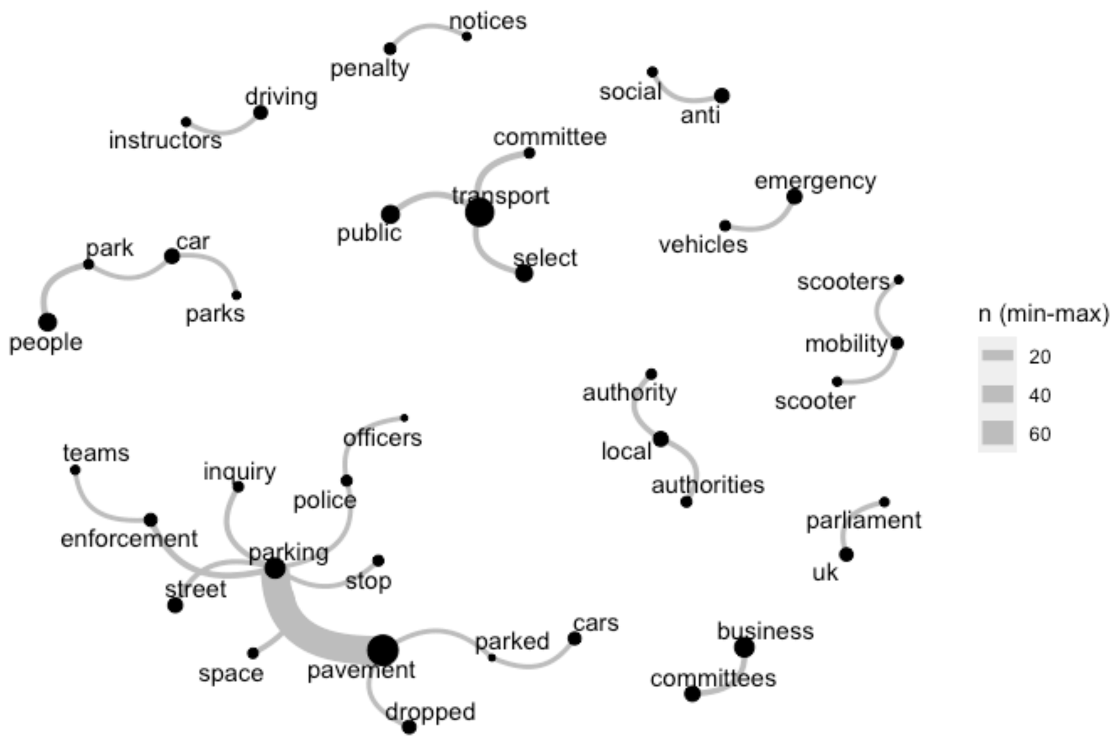
As the number of topics in this discussion was only three, I used the highest tf.idf measures for words in each topic to have a clearer understanding of which issues are raised. Figure 70 shows the three topics (created and named by the committee staff) and reveals the "Enforcement" topic is focussed on financial aspects of pavement parking such as "revenue", "budget", "income", and "pay", as well as effects on local areas like "rural" and "neighbourhood". Topic "Sanctions" is concentrated on effects of pavement parking on the public such as congestion, risk, access and bump. There is also discussion about the legality of parking on the pavement with words such as "crime" and "court" suggesting participants were in favour of criminalising the practice. Finally, the "Solutions" topic had a clear dominating word, "ban", suggesting participants favoured prohibiting pavement parking as the best solution.

Figure 70: Transport Committee (Pavement Parking) Discourse seed topic word distribution ranked by tf.idf measure



A network of bigrams (Figure 71) within the Discourse discussion revealed "pavement parking" and "parking wardens" as the most frequent bigrams. A cluster of bigrams comprising of phrases such as "mobility scooters" and "wheelchair users" related to the Enforcement and Sanctions topics and point to issues for people who may be more affected by pavement parking than others, while another cluster of phrases surrounding "penalty notices", "enforcement teams", and "anti social" suggest possible remedies to the problem of "anti-social" pavement parking and closely aligns with the Sanctions topic. The bigram cluster on "parliament inquiries/business/news" shows the participants taking an interest into the wider business of the committee and perhaps following the inquiry's updates after the discussion.

Figure 71: Transport Committee (Pavement Parking) Discourse bigram network



Looking at the sentiment expressed in this discussion on Discourse highlights that many comments were displaying sentiments of *fear, surprise* and *sadness* such as "there is no need for pavement parking unless you are on a very narrow street. Parking wardens get terrible abuse by the public so I believe it should be enforced by Police back up. It is a Constant problem that is usually totally unnecessary to park on pavements putting partially sighted/blind people, disabled and families with children into a dangerous situation" contributing most to these sentiments.

Table 16: Transport Committee Twitter - LDA statistical validation metrics

topics	Griffiths2004	CaoJuan2009	Arun2010
10	-15397.77	0.22932854	442.6978
9	-15501.81	0.15358416	449.8091
8	-15575.67	0.14065104	460.8343
7	-15597.2	0.17059249	469.5035
6	-15727.78	0.24265276	490.7451
5	-15851.84	0.2067811	509.9433
4	-16017.88	0.06482446	515.628
3	**-16399.62**	**0.04838395**	**544.0328**

2	-17019.83	0.22516321	597.7565

Figure 72: Transport Committee (Pavement Parking) Twitter LDA topics word distribution ranked by tf.idf measure

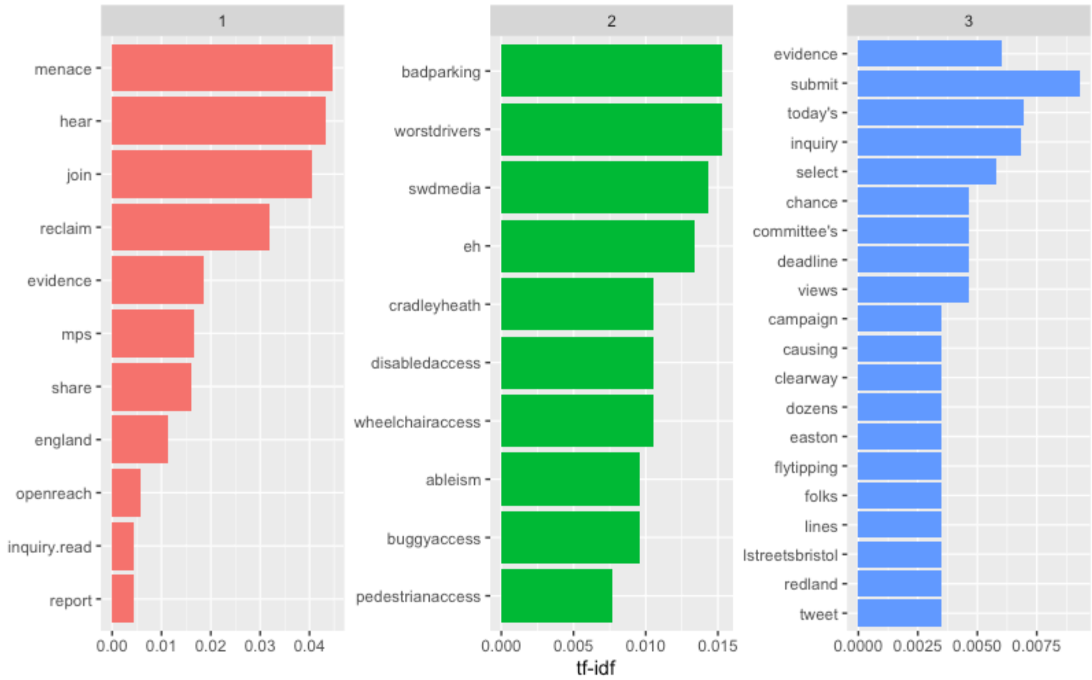


Figure 72 shows the Transport Twitter discussion using the hashtag #PavementParking segmented into three topics using LDA which had a low alpha value of 0.030. The statistical validation metrics for the LDA model are displayed in Table 16, and show that 3 topics provides a good balance between minimising the Arun2010 and CaoJuann2009 metrics and maximising the Griffiths2004 metric. The first topic features words such as “menace”, “hear” and “reclaim” as words with the highest tf.idf scores suggesting this topic was more concerned with expressing frustration with the pavement parking practice. This topic also contains words such as “mps”, “evidence” and “report” which also points towards a focus on the committee’s inquiry and the progress of the report. The second topic appears to highlight the hashtags which were used in this discussion such as “badparking”, “worstdriver” and “disabled access” also suggesting a range of online discussions surrounding the same issue were being included into this conversation. Similarly, to topic 1, topic 3 was focussed on the work of the select committee with words such as “submit”, “inquiry”, and “evidence” notifying people of the committee’s upcoming inquiry report.

A large cluster of bigrams (Figure 73) surrounding “pavementparking” led to different feelings towards the issue being expressed such as “illegal parking”, “selfish”, and “worstdrivers”. Many of these words are also identified in the negative sentiment classification in Twitter. Possible negative implications of pavement parking are also picked up in the bigram network with phrases such as “disabledaccess”, “flytipping”, and “ableism” identifying specifically people in wheelchairs affected by pavement parking. Smaller clusters with “dual carriageway”, and “primary school” highlight different areas of the community which are affected and shows the range of topics which were being discussed using the

#PavementParking hashtag during the week. The majority of tweets in the Twitter discussion expressed *positive, negative, anticipation* and *trust* sentiments for example "More #PavementParking in #cradleyheath to avoid the £2 car park charge probably!#selfishparking #illegalparking #pedestrianaccess #wheelchairaccess #buggyaccess #disabledaccess #ableism #ignorant #selfish https://t.co/dORmtSsZyj" and "A Colchester councillor explains why Traffic Regulation Orders aren't a magical solution to #PavementParking problems, and why they want wider powers for the Council: https://t.co/ZzLq9ovURG".

The analysis shows a similarity in issues raised between the discussions on Discourse and on Twitter, especially regarding identification of those most affected by pavement parking. However, these bigram networks also show differences in the subject matter across the two platforms. There was more of a focus surrounding solutions and measures to manage pavement parking on Discourse than on Twitter. The reasons for this difference in focus between the two platforms could be various. Twitter has much less ability to have deliberative, coherent discussions which in this context could lead to an absence of discussions surrounding how a problem can be resolved. A solution-driven discussion, as the one seen on Discourse, not least because Solutions was a seeded topic in this coordinated discussion, would require an understanding of what the problem is, but the next step would be assessing how the problem can be resolved. I see that the Twitter discussion had a larger focus on the theme of identifying problems and blaming people for pavement parking rather than exploring possible solutions to the issue. Combining this analysis with the sentiment analysis from the two channels (Figure 74) shows that this emphasis on problems and solutions on Discourse also results in an emotionally more varied discussion on the platform.

Figure 73: Transport Committee (Pavement Parking) Twitter bigram network

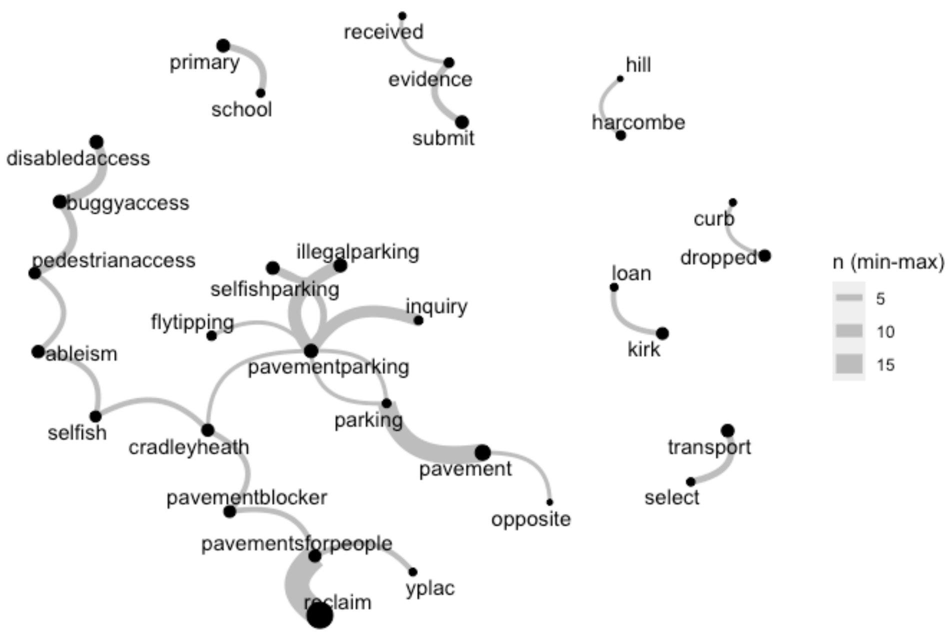
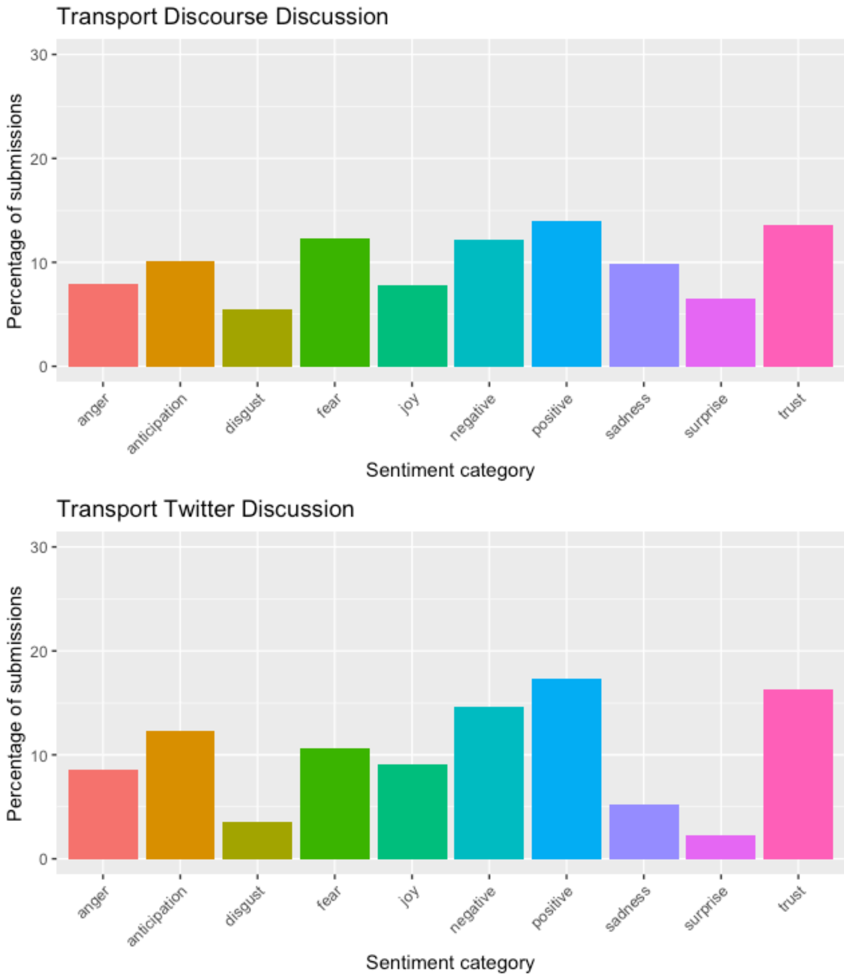


Figure 74: Transport Committee (Pavement Parking) Dicourse and Twitter sentiment distribution

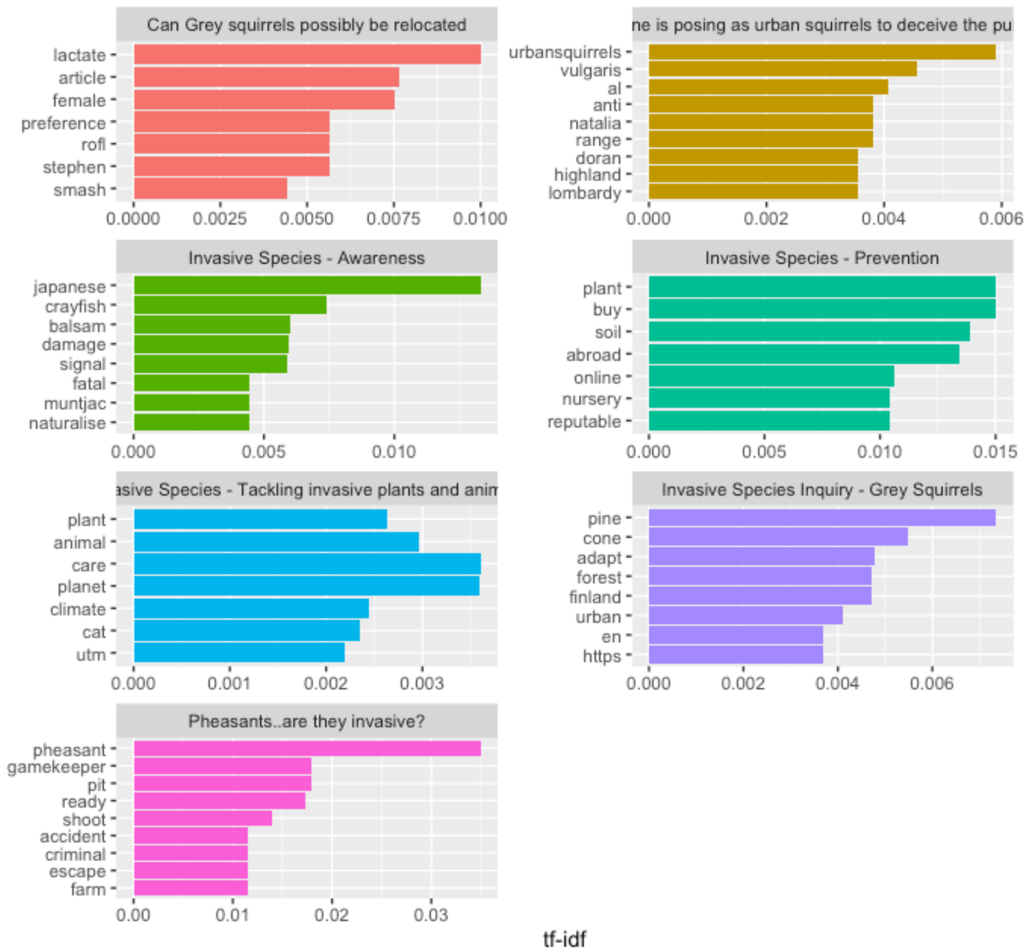


6.6.2.2 Environmental Audit Committee 'Invasive Species'

The Invasive Species discussion held by the EAC committee had 7 topics on the Discourse platform, also a manageable number to analyse without the use of topic optimisation and modelling. Instead, I analysed the words with highest tf.idf scores in each of the Discourse topics. Topics "Awareness" and "tackling invasive plants and animals" are primarily focussed on discussions about invasive species such as Japanese knotweed and Himalayan balsam, climate change as one of the causes for the invasive species problems as well as the effects of invasive species on ecosystems they invade. This latter topic also featured the most comments out of all with 240 responses. The "Prevention" topic is more concerned with methods of prevention of invasive species discussing issues with buying soil abroad, buying reputable seeds, and potential quarantines for goods imported. The remaining topics "Can Grey squirrels possibly be relocated", "Fraud alert - someone is posing as urban squirrels to deceive the public and discredit us", and "Grey Squirrels" all focus on the topic of squirrels, specifically the Eurasian/Vulgaris, i.e. red squirrels which is threatened in its habitat by the invasive species of grey squirrels. Discussed in this context is euthanasia as a barbaric method of killing grey squirrels in order to protect the red squirrels. Furthermore, it acknowledges that the grey squirrel is better in adapting to different environments. Finally, the topic “Pheasants..are they invasive?” was focussed on discussing the categorisation of other wildlife as invasive species with words such as “escape” and “farm” suggesting that these invasive species often escape and then become a nuisance.

There were also a few mentions of participants’ names who were heavily involved in the discussion. For example, topic “Fraud alert - someone is posing as urban squirrels to deceive the public and discredit us” includes ‘natalia”, topic “Can Grey squirrels possibly be relocated” has “Stephen”, and topic “Grey Squirrels” has “mohutchinson” who can all also be

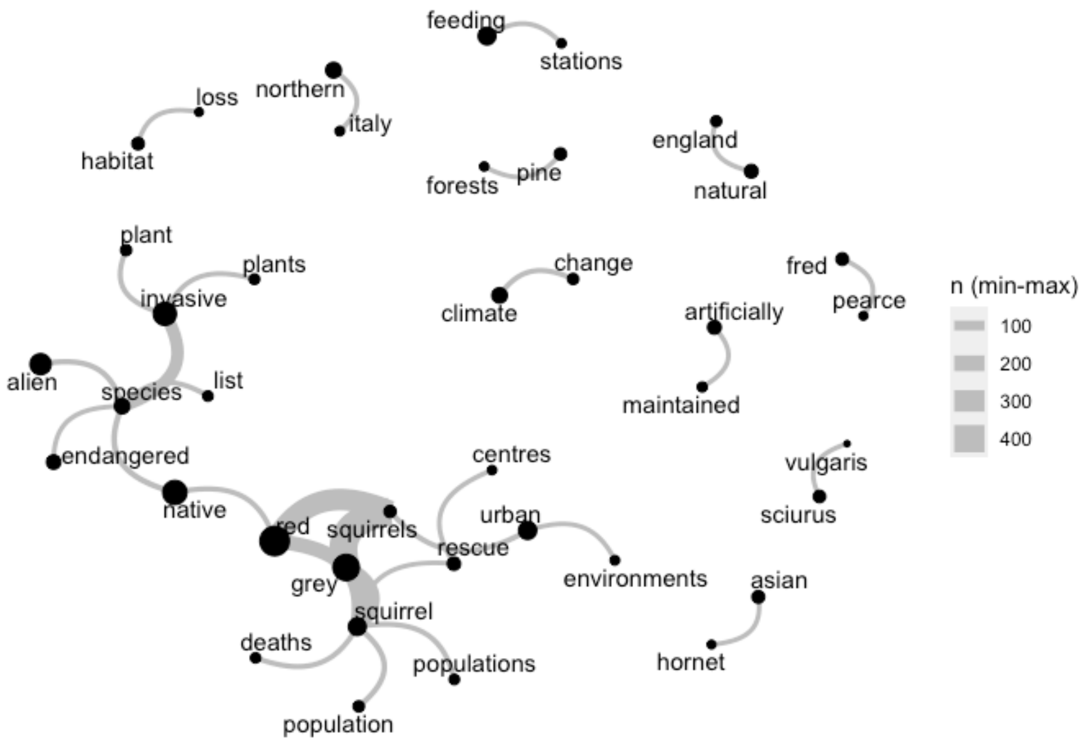
Figure 75: EAC Committee (Invasive Species) Discourse user defined topics word distribution ranked by tf.idf measure



found in the Discourse network of replies shown in Figure 64.

Examining now the bigram-based network, I can see that “grey squirrels”, “red squirrels”, and “invasive species” are the most common word pairs in the Discourse dataset (Figure 82). A cluster comprising of phrases such as “alien/invasive/native species” show a difference of terminology relating to invasive species used in the discussion, while another cluster of phrases surrounding squirrels refers to a range of topics which focussed on the treatment of grey squirrels in order to protect red squirrels. “Japanese knotweed”, “asian hornet” and “climate change” are also common bigrams and suggest other discussions that were ongoing (relating to topics “Awareness” and “tackling invasive plants and animals”) independent of the squirrel debate. Participants in this Discourse discussion expressed sentiments of *trust, fear,* and *sadness*. For example a comment ranked highly in the *trust* sentiment category was “I now urge the goverment to show compassion , take scientific facts into account not myths or lies and make UK grey Squirrel rescue and release exempt from the Invasive species order. People have had enough of wrongly blaming Grey’s for any red decline. Grey’s are an important part of Brittish wildlfe and People’s wellbeing . Compassion is not a crime”. This comment expresses a trust in the UK government to take action to ensure grey squirrels are not categorised as an invasive species. Likewise, a comment categorised as *fear* reads “Tackling an invasive plant is far different to a so-called invasive animal. One has a central nervous system, is capable of feeling fear, pain and compassion. It’s humans that are the ones who are invasive and destructive. I destroyed, and continue to ruin the space my native animals need. I brought the grey squirrels here and now because I deem them to be a pain I decide to kill them off?! It’s unfortunately my answer to everything these days.” Here, the fear is using a generalised method of categorising plants and animals as invasive species and the difficulties that raises for native species.

Figure 76: EAC Committee (Invasive Species) Discourse bigram network

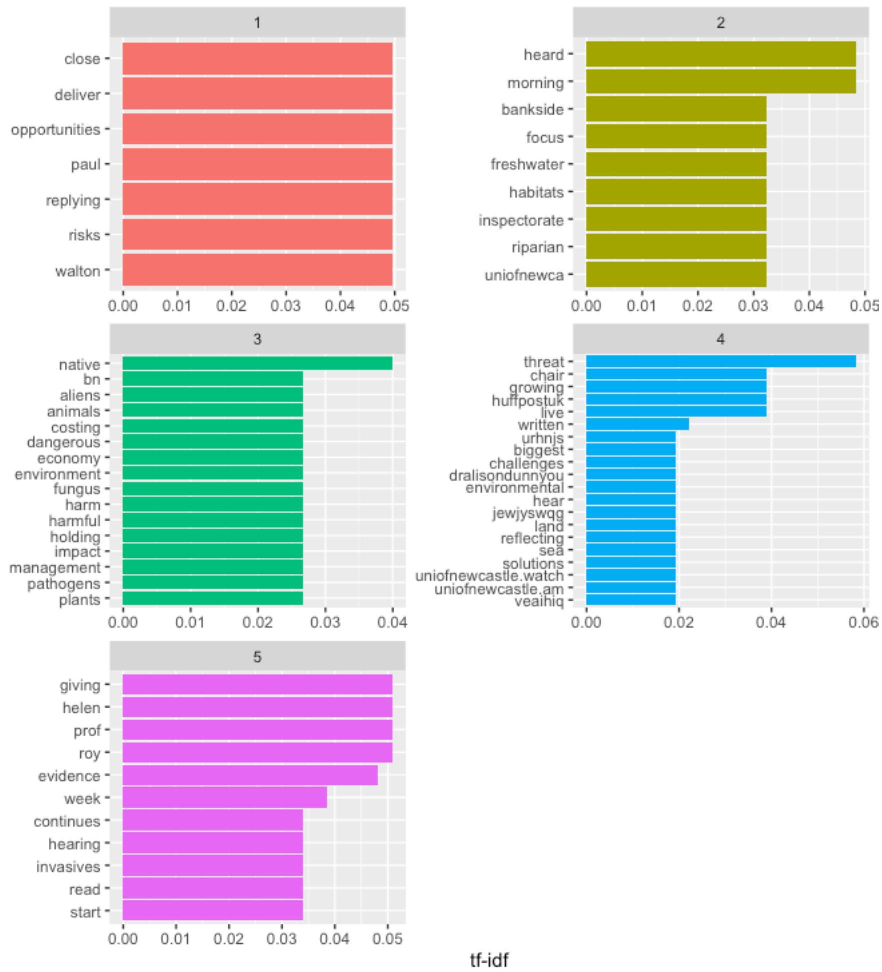


Using LDA, the topics for the Twitter discussion are shown below. LDA validation in Table 17 revealed the optimal number of topics as 5 as this minimised the Arun2010 and CaoJuann2009 metrics and maximised the Griffiths2004 metric. This model had a low alpha value of 0.018 suggesting comments contain a mixture of only a few topics and are therefore relatively different to each other. The words with the highest tf.idf scores are plotted in Figure 77. This shows that topic 1 does not provide very coherent topic structure with words such as “close”, “deliver”, and “opportunities” featuring highly in this topic. Topic 2 is focussed on invasive species affecting environments around water with words such as “bankside”, “fresh water” and “habitats”. Topic 3 focusses on the harm invasive species are doing to different environments specifically the “pathogens”, “fungus” and the negative impact on the "economy". Topic 4 concentrates on the "growing" "threat" of invasive species and highlights the University of Newcastle as a potential for "solutions". Finally, topic 5 again highlights leaders in the field such as "helen roy" and the process of "giving" "evidence" to the committee's inquiry.

Table 17: EAC Committee: LDA statistical validation metrics

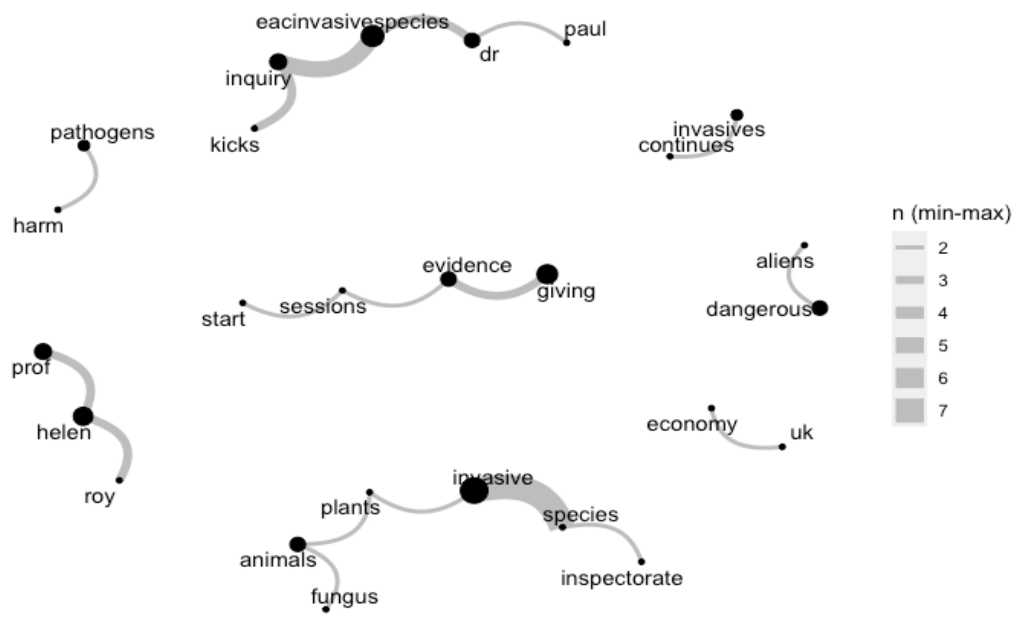
topics	Griffiths2004	CaoJuan2009	Arun2010
10	-1601.268	0.171	25.155
9	-1606.737	0.154	26.119
8	-1608.206	0.154	27.980
7	-1611.495	0.152	30.116
6	-1617.266	0.156	32.925
5	**-1655.274**	**0.115**	**34.851**
4	-1664.677	0.128	39.193
3	-1699.322	0.134	44.318
2	-1767.294	0.110	51.754

Figure 77: EAC Committee (Invasive Species) Twitter LDA topics word distribution ranked by tf.idf measure



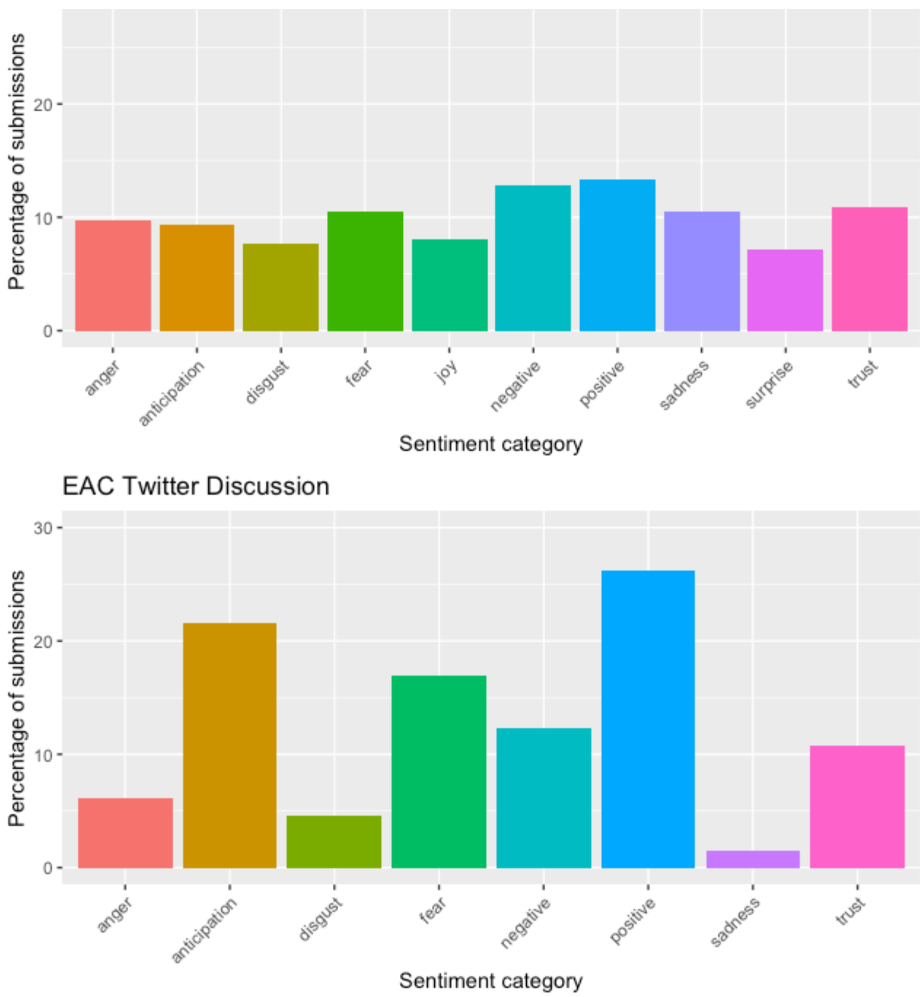
The Twitter discussion bigram network (Figure 78) also shows there were repeated mentions of ecologist "prof helen roy" owner of @ukladybirds account and co-chair of an IPBES Invasive alien species study, and discussions about the evidence sessions held by the Environmental Audit committee. Issues surrounding different invasive species were raised on Twitter identifying plants, animals, and fungus, invasive species (aligning with topics 2 and 4) and expressing concerns such as "growing threat", "uk economy 2bn", and "harm pathogens" (aligning with topic 3). The bigram-based network indeed shows the range of topics which were being discussed using the #EACInvasiveSpecies hashtag during the week. The Twitter discussions were also held during Invasive Species Week on Twitter where many organisations were holding events, but Twitter event used a separate hashtag #InvasiveSpecies. The EAC Twitter discussion was primarily positive with tweets about continuing to follow the committee's inquiry "#InvasiveSpecies week is over but my work on invasives continues. Today we are giving evidence to the @CommonsEAC. Read the #EACInvasiveSpecies written submissions here https://t.co/kuXocZoda9 https://t.co/lLGsgLYD9U". However, there were also feelings of anticipation for the evidence session being held for the inquiry "Our first evidence sessions start next week. We're hearing from Prof Helen Roy @UKLadybirds, @WCL_News, @theCCCuk, @UniofNewcastle, @UniOfYork, @WayneDawsonEco, @ThinkUHI, and @DrAlisonDunn. Follow my inquiry at https://t.co/0YFi1Mfr2l and #EACInvasiveSpecies @InvasiveSp"

Figure 78: EAC Committee (Invasive Species) Twitter bigram network



The EAC 'Invasive Species' discussion had the largest difference in sentiment between the two platforms (Figure 79). The Discourse discussion was mostly negative with words such as "killing" and "damage" being used frequently, while one in ten of the comments in this discussion were flagged as inappropriate by other participants. While the Discourse discussion had a fairly uniform distribution of all sentiments, the Twitter discussion featured a high proportion of *positive* and *anticipation* sentiments making up 47% of all comments. The word 'inquiry' featured heavily in these two sentiments with comments such as "#InvasiveSpecies week is over but my work on invasives continues. Today we are giving evidence to the @CommonsEAC. Read the #EACInvasiveSpecies written submissions here https://t.co/kuXocZoda9 https://t.co/lLGsgLYD9U". This committee's discussion was also more varied on Discourse than Twitter, with a range of sub-topics covered which could explain the varied sentiments. For example, there was a large discussion surrounding the categorisation of red and grey squirrels on Discourse, but this did not occur on Twitter. Furthermore, the ability to create sub-topics on Discourse naturally allowed participants to organise their discussions based on the different issues they had an interest in, and allowed the committee to raise specific issues they wanted views on using the seed topics. The lack of this feature on Twitter meant that users were unable to go in depth and focus on a range of sub issues. On Twitter, the discussion was more focussed on a small set of sub-topics, namely the process of giving evidence to the committee, several key users in the discussion, and several species which are considered invasive in the UK.

Figure 79: EAC Committee 'Invasive Species' Dicourse and Twitter sentiment distribution



6.6.2.3 EFRA Committee ‘Plastic Packaging’

Finally, the EFRA discussion on plastic food and drink packaging had 142 topics, which were condensed into 12 topics using the LDA tuning package for easier interpretability. This model also had a low alpha value of 0.0454 suggesting comments contain a mixture of only a few topics and are therefore relatively different to each other. Since this Discourse discussion differed from the other committees’ discussions in that it required topic modelling of the many user-generated sub-topics, I used the posterior probabilities of the LDA model trained on the Discourse comments, and applied them to the tweets. In this way, I could observe how the discussions on both platforms differed with respect to the same topics. Therefore, the word distribution in each of the 12 topics remains the same but the number of submissions which align with those topics changes between platforms. Distribution of the topics on Twitter and Discourse are found in Figure 80 and summaries of the words in these topics are displayed in Figure 81 and

Table 18.

Figure 80: EFRA Committee (Plastic Packaging) LDA topic distributions

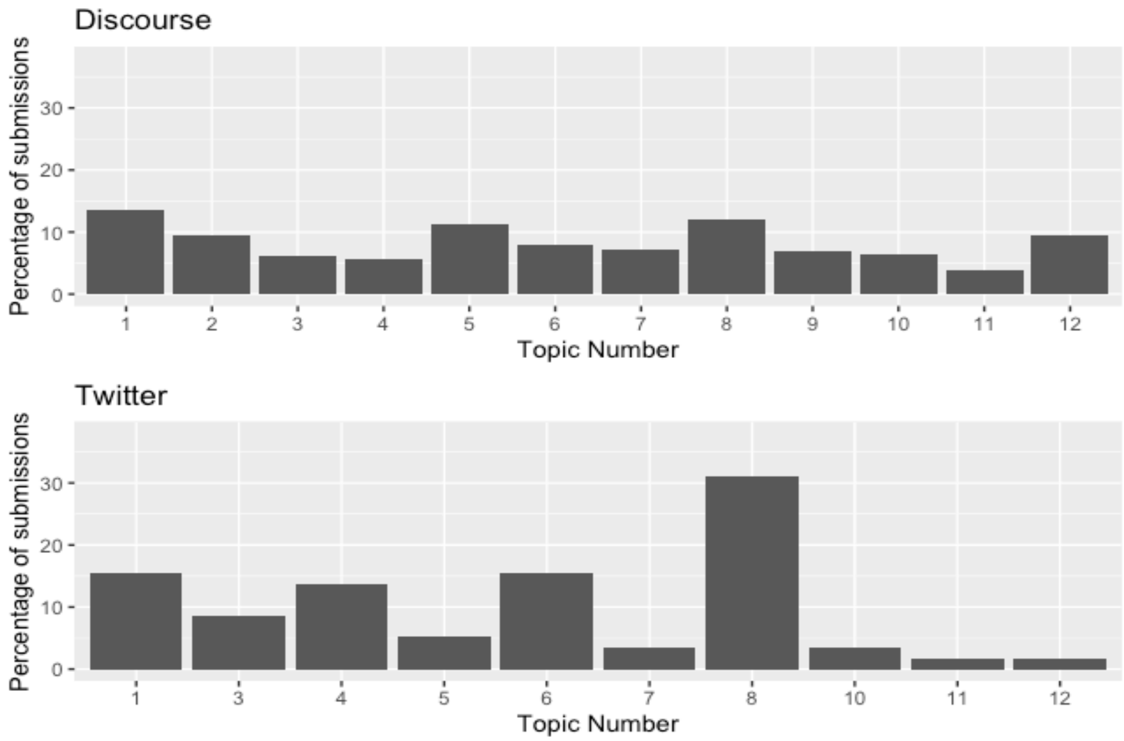


Figure 81: EFRA Committee (Plastic Packaging) LDA topics word distribution ranked by tf.idf measure

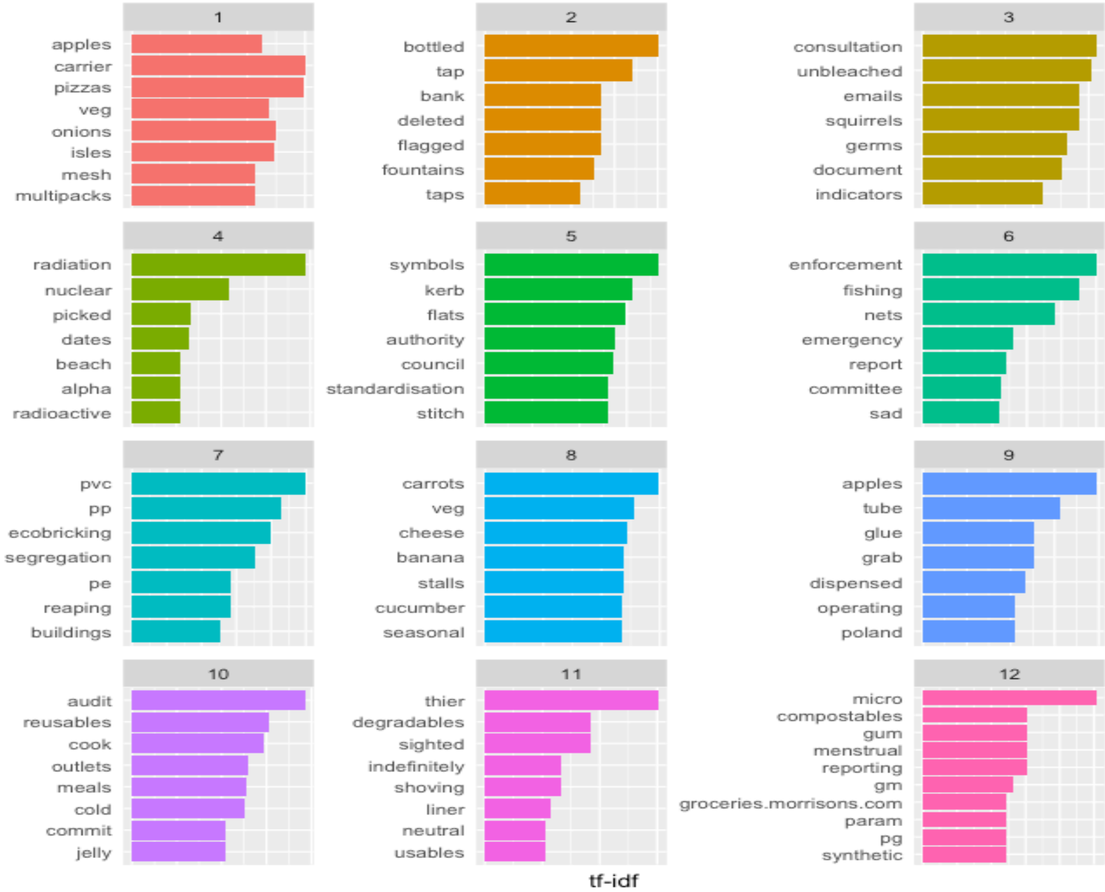
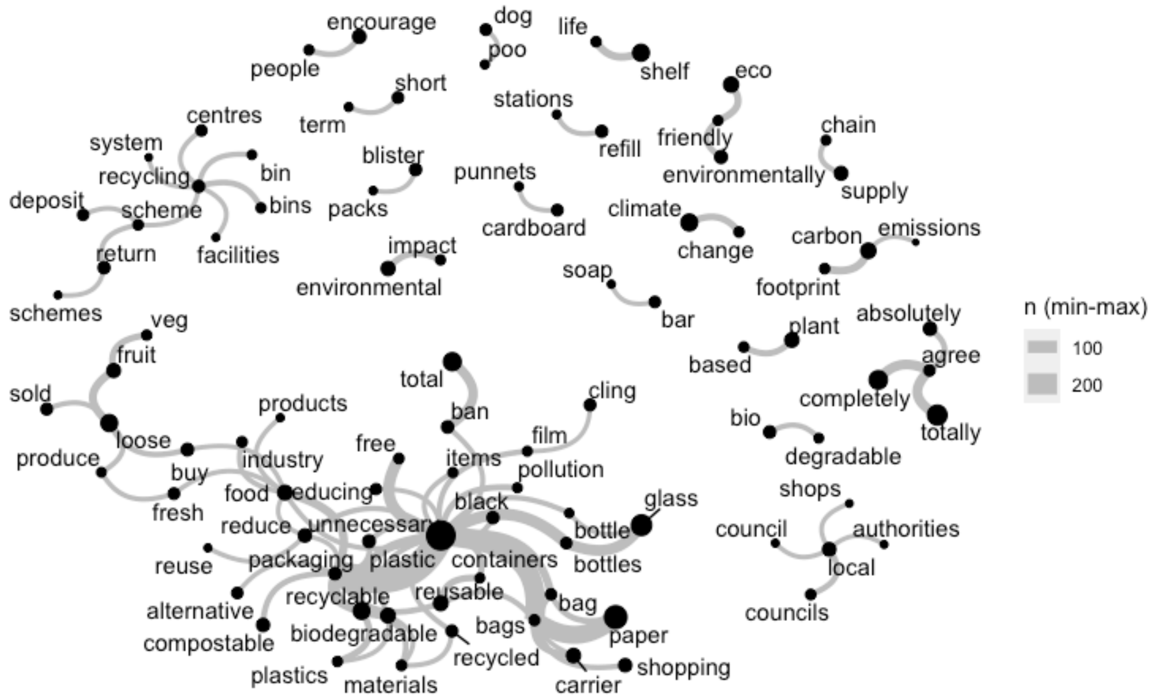


Table 18: EFRA Committee (Plastic Packaging) - Summary of LDA topics

Topic 1	packaging for fruit and veg
Topic 2	water
Topic 3	Squirrels, details about the inquiry and submissions
Topic 4	Problems caused by radiation
Topic 5	Local government action to combat wasteful packaging
Topic 6	Fishing enforcement, mentions of the committee report
Topic 7	Different types of sustainable practices and materials
Topic 8	Seasonal fruit and veg
Topic 9	glue tubes
Topic 10	audit of the food industry
Topic 11	Different packaging materials
Topic 12	other non-recyclable products such as menstrual products and gum

This was a very popular discussion on Discourse where many different areas of discussion were raised over the week (Figure 82). "Recycling return scheme/s", "recycling bins", and "recycling centres" were mentioned often in the bigram network relating to topics 5 and 7. "Local council/authority/shops" were also mentioned in relation to the "food supply chain", "fast food" and "fresh products". This also related to a discussion about "loose fruit and veg" and encouraging people to buy these instead of those packaged in plastic as in topic 1. Discussions about "home composting", "reusable coffee cups" and "reusable fizzy drinks containers" were also raised along with calls for "tax incentives" to "reduce plastic". "Biodegradable materials" and "glass bottles" for milk and "drinking water" are shown as possible solutions. Finally phrases such as "environmental cost", "future generations", and "climate emergency" show a particular concern of the participants surrounding the lasting damage that is being caused. The presence of "squirrels" in topic 3 suggests a potential overlap of users involved in the previous discussion of invasive species from EAC.

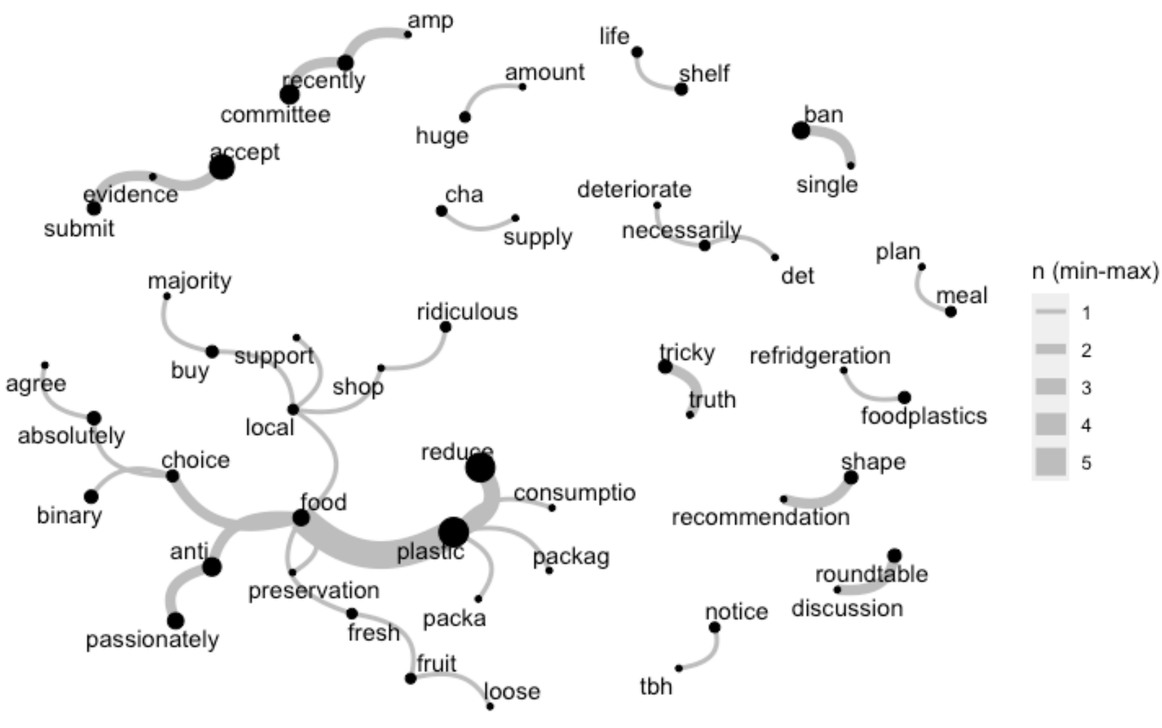
Figure 82: EFRA Committee (Plastic Packaging) Discourse discussion bigram network



Regarding the sentiments expressed, the Discourse discussion had more *trust, anticipation, anger, disgust,* and *fear* sentiments in the comments however none of these were considerably more prominent than the other. An example of a comment categorised as both *trust* and *anticipation* is "We have the technology to allow customers to buy food with next to no packaging, to weigh their food and place it in their biodegradable or reusable packaging and then add it to their shopping trolley while logging the content on a self serve shopping device, why don't we start using it". This comment could be defined as *anticipation* as it is encouraging the people to use the technology in place to make their shopping more environmentally friendly. However, it appears to have been mislabelled also into the *trust* category. The comment "Producers will only change if government and / or the public enforces action. i.e …bans by the government and refusal by consumers to purchase single use plastic If the producers, in a competitive market i.e supermarkets , find they are losing sales because the packaging is either disapproved of by the purchasers , and / or the government market forces could lead to change. However, we are in need of URGENT solutions to plastic waste that can only be introduced by government intervention.", is categorised into the *fear* and *negative* sentiments. This time, the categorisation of the NRC lexicon appears to be accurate with the user expressing concern of the time sensitive nature and need to introduce interventions quickly.

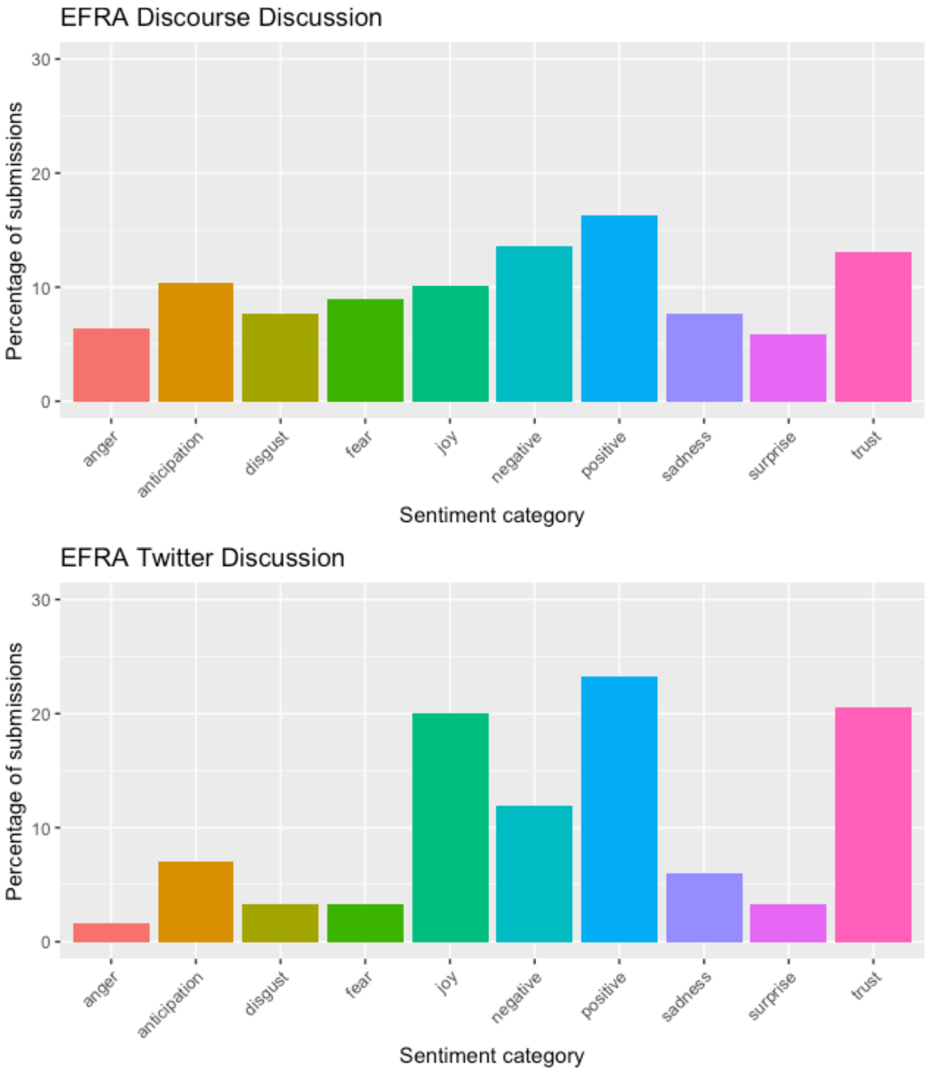
The Twitter discussion using #FoodPlastics revealed phrases such as "anti food plastic packaging" and a requirement for choices in food shopping and packaging (Figure 83). This concern about food waste and packaging is also shown as a prominent issue in the topic distribution. There were also concerns surrounding "foodplastics refrigeration", "tricky truth", and "supply chain". Several users spoke of more procedural matters to do with the committee and the inquiry, with phrases such as "submit evidence", and "shape recommendations" so there was a clear awareness that this Twitter discussion was being held by Parliament. The large presence of these committee-related phrases appear exclusively in the Twitter bigram networks of the discussions, especially in the case of EFRA and EAC (Transport less so), suggesting the overt presence of the committee accounts facilitated a connection between them and the Twitter users engaging with the topic.

Figure 83: EFRA Committee #FoodPlastics Twitter bigram network



This Twitter discussion again featured a small set of prominent sentiments, primarily *positive, joy,* and *trust*. These sentiments accounted for 63% of the tweets on this platform (Figure 84). Words such as 'food', 'committee', and 'inquiry' featured heavily including comments such as "Truth is #plastic often used to create an illusion of freshness. Makes old food look better longer." and "Food preservation or waste need not be a binary choice. Absolutely agree about reducing plastic packaging and we need biodegradable plastic alternatives such as paper and wax derivatives". The topic distributions (Figure 80) also show that while Discourse was not heavily dominated by any single topic, Twitter was much more focussed on seasonal fruit and vegetables (topic 8) and how they are packaged. Therefore, the concentration on a single topic in a discussion can also cause the sentiment of that discussion to be heavily biased.

Figure 84: EFRA Committee (Plastic Packaging) Discourse and Twitter sentiment distributions



Conclusion

The first half of this chapter has analysed several online engagement activities held by the UK Parliament throughout 2018. Each section focussed on a particular research question (section 1.1.2) and area of interest for the teams analysing these discussions and sought to emulate the summaries they would usually make manually. Each discussion in my sample exhibited different characteristics based on the type of people, topics, and sentiments expressed in each. Analysis of the sentiments in the Animal Fur discussion show it to be overwhelmingly negative, more so than any other discussion in the sample. It is possible that rather than encourage useful discussions, this discussion and the erratic temporal pattern of posting may have contributed to its negative sentiments. Unlike the other discussions, Visit Visa survey had several main topics that were prominent in the survey rather than one single topic dominating the comments. This could be a result of the design of the survey which was split into three questions, or could be down to the lack of interaction participants had with each other. As they were unable to see each other's comments, the submissions were not dominated by one single issue that users see and respond to. Rather, the participants in the Visit Visa survey had the opportunity to speak about a range of sub-topics that were important to them without the influence of other participants' views and opinions. Therefore, this format of a survey could facilitate the emergence of new ideas easier than a discussion on Twitter or Facebook.

The majority of analysis in this chapter has explored each discussion as a whole, however by breaking a discussion down by socio-demographic clusters of its participants, I find that users on the extremes of the age indicators put more emphasis on a select few sub-topics than users using more general language which did not provide much indication into their age. The second half of this chapter has evaluated the demonstration tests conducted between May and June 2019 with several UK Parliament select committees. I have experimented with the use of existing online platforms used by the UK Parliament such as Facebook, Twitter and surveys, while introducing a novel platform, Discourse, which provides an alternative approach to engagement. Discourse prioritises users interacting with each other and spending enough time on the platform through their trust system which in turn fosters balanced discussions.

By combining the analysis of geographic spread, subject matter, sentiment, and network characteristics, I can build an insightful picture into the nature of these discussions across the two platforms. The location of the participants is heavily concentrated to England and there were not many Scottish, Northern Irish, or Welsh citizens participating with any of the three discussions on Discourse. Discourse discussions were much more connected in terms of social network interactions and much more nuanced in terms of sentiments expressed and topics discussed, with more space to cover different sub-themes within the same issue and with a clearer commitment to identifying solutions to raised issues. Twitter had a less connected discussion in terms of social network interactions, however there was evidence of greater clustering around central hub-accounts. The themes discussed on Twitter were less nuanced and less solution-focused across all three committee discussions. Gathering everything I have learned in these demonstration tests suggests Discourse is the better platform for online engagement. It allows for the creation of specific topics which allow participants and administrators to guide the discussion. This gives the discussion some structure and guidance for the many different sub-topics that can occur in an engagement session, while providing participants with a specific outlet for them to voice their opinions. The social network graphs from Discourse and Twitter also show that users in the Discourse discussions were generally more likely to engage with each other during the discussion than the Twitter users. This creates more focussed and rich conversations which address multiple aspects of a discussion which a select committee team is interested in. This also saves time and resources, as to address each sub-topic separately on Twitter would require separate hashtags and more maintenance for committee staff.

A benefit of using a purpose-built platform such as Discourse was its moderation feature which was valuable to both participants and staff. It allowed participants to take action against suspected trolls or those who they felt were being inappropriate, and was used mostly during the EAC discussion surrounding red and grey squirrels (section 6.6.2). Social media platforms such as Twitter do not have this type of moderation built into their structure which can cause off-topic discussions to overwhelm an online debate with little capacity from the institution to get it back on track. Had the same EAC discussion about squirrels appeared on Twitter instead of Discourse, participants would not have had the ability to create a separate topic for this, and committee staff would not have been able to address any inappropriate or flagged comments.

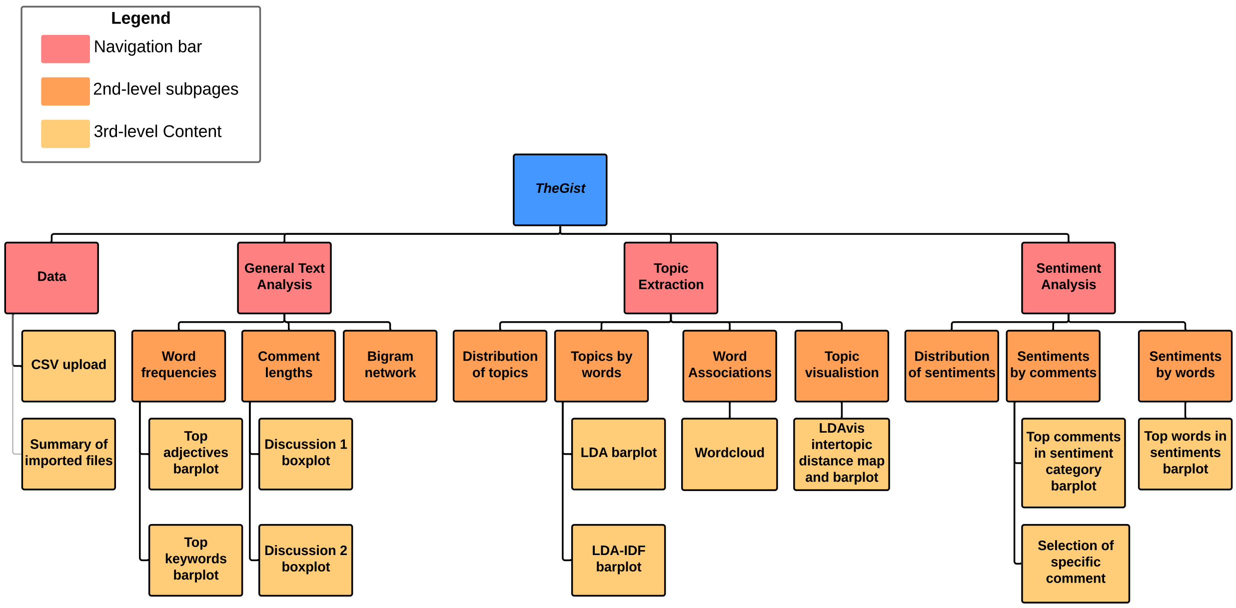
The interactive reports produced as a result of the demonstration tests have been used by officials across Parliament in committee reports and for questioning of ministers and could make a valuable contribution to further policies and strategies. The final research question of this PhD is 'how can citizen input be utilised in a meaningful way to inform policy making?'. This chapter has provided an answer which does not negatively impact the daily workings of committee staff, but reduced their final workload, allowing them to easily harness citizen input using text mining without losing any of the public's voice.

Chapter 7 TheGist: A new analysis tool for parliaments

Throughout this project, text mining has been paramount to understanding how citizens feel about the issues they are discussing in with online debates. However, at present this analysis can only be achieved by those with the necessary data analysis and natural language processing skills. For staff in the Digital Engagement team, select committee teams and other departments who work with digital engagement in Parliament, learning these skills is not a priority and would take valuable resources away from their day-to-day work. However, through the course of this project, parliamentary staff have learned to appreciate the insights gained through this type of analysis[41].

Therefore, to facilitate the analysis and evaluation of online discussion comments in the future, I decided to create a web application including many of the methods described in Chapter 3. This application is called TheGist[42] and this section explains how it works. TheGist was created using R statistical programming software, specifically the RShiny package (RStudio Inc., 2013) which allows users to create dashboards, user-interfaces and web applications. Visualisations integrated within TheGist were created using the ggplot package (Wickham, Chang and Wickham, 2016) and NLP models using the tm, tidytext (Silge and Robinson, 2016), quanteda, topic models (Hornik and Grün, 2011), and udpipe (Wijffels, 2019) packages. I adapted the appearance of the application frontend to closely reflect the House of Commons branding using CSS.

Figure 85: TheGist application map



As shown in Figure 85, TheGist is divided into 4 sections: Data, General Text Analysis, Topic Extraction, and Sentiment Analysis. These sections each address a specific method in natural language processing which has been used throughout this project and shown to add value to the analysis and understanding of parliamentary engagement activities.

---

41 Personal Communication, Westminster

42 https://github.com/NicoleDNisbett/TheGist and https://nicolednisbett.shinyapps.io/TheGistDemo/

It was important for the tool to be made with the end user in mind and be as specific to their needs as possible. The Digital Engagement team and other teams in the UK Parliament had previously used a social media listening tool to analyse their social media channels' activities however this was abandoned after a short time. Through speaking with several teams across Parliament who used that tool, it became apparent that it had too many unhelpful features and was difficult to learn (more details in section 4.3). To avoid this problem occurring again with TheGist, each method of analysis and visualisation is carefully chosen to provide the user with a tool which is useful but not overbearing.

7.1 Data

This section of the application works as a landing page to explain how TheGist works, upload data and view which files are already in the system. Currently it is possible to upload csv files. Data from Facebook in csv format can be for instance obtained through Socialfy.pw, which is a free application allowing owners of a Facebook page to download comments made in response to a specific post (Socialfy.pw, 2020). This download captures the names of users who interacted with the post, their comment text, a timestamp, and information on how many likes the comment received. For the purposes of TheGist, only the comment text is used. If the raw data is in another format or from another source, it can still be imported into TheGist using this csv tab as long as it contains a column with the comment text.

When importing a file into the application, a summary of the file contents is displayed so the user can confirm they have the correct information before the import begins. Once happy, the user selects the 'save file to app' button which fully imports the file into TheGist and prompts the user to type a unique name to refer to the file in the application. This feature allows the user to change the name of the file to something more memorable and useful than the automatic filenames which are assigned by the Socialfy.pw download for example. Once this file is uploaded, its details are included in a summary table. This table contains the variables File Name, Assigned Name, and Number of Comments, and summarises all files which have been imported into the application. The Assigned Name variable shows which unique name the user had selected for a file, and is used in the rest of the application to refer to the correct file. This summary table also ensures the same file is not uploaded multiple times to the application.

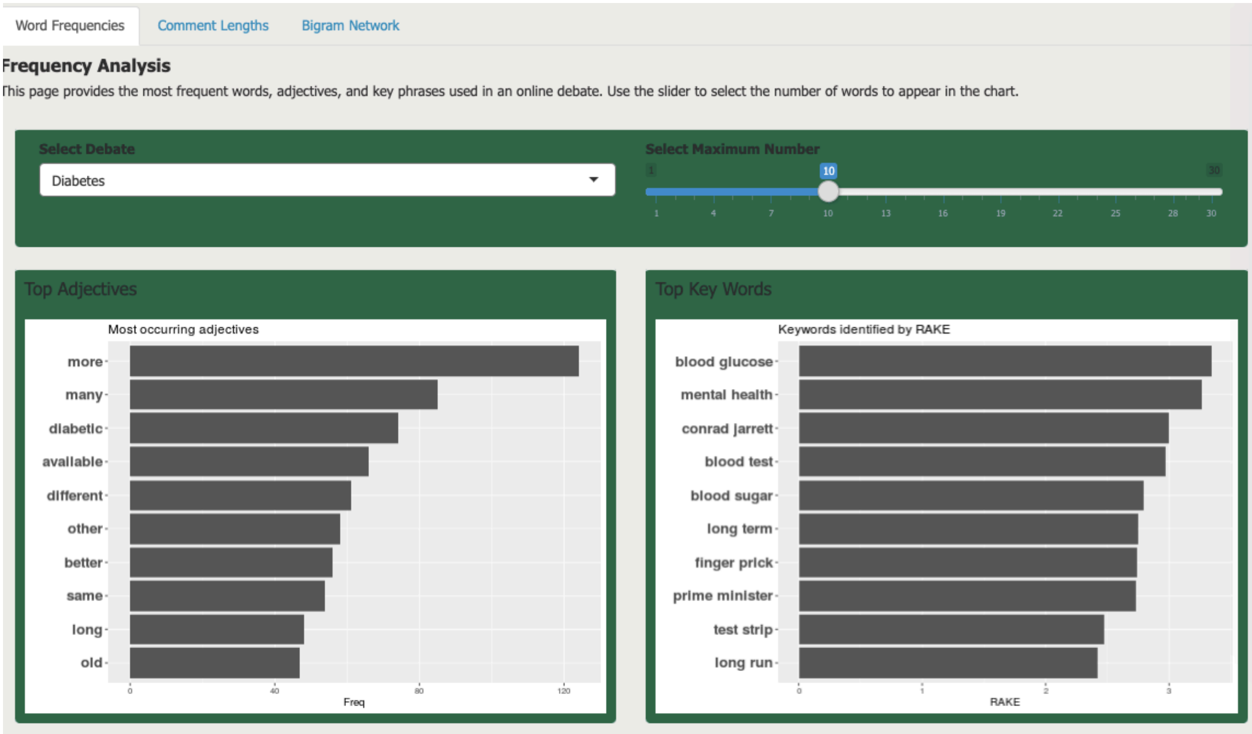
7.2 General Text Analysis

The General Text Analysis section of the application is intended to give the user a summary of the comments in the files they have uploaded. It includes three subpages: Word Frequencies, Comment Lengths, and Bigram Network.

7.2.1 Word Frequencies

This subpage allows the user to select an imported file using a dropdown list, and displays two barplots. The first displays the most frequent adjectives used in the discussion, and the second shows the top keywords in the discussion as identified by the Rapid Automatic Keyword Extraction (RAKE) model (Rose *et al.*, 2010). The number of words displayed in these two plots can be altered using a slider. The most frequent adjectives and keywords are obtained using part-of-speech (POS) tags of the comments and word co-occurrences to extract the main word phrases occurring in a text.

Figure 86: TheGist Word Frequencies snapshot



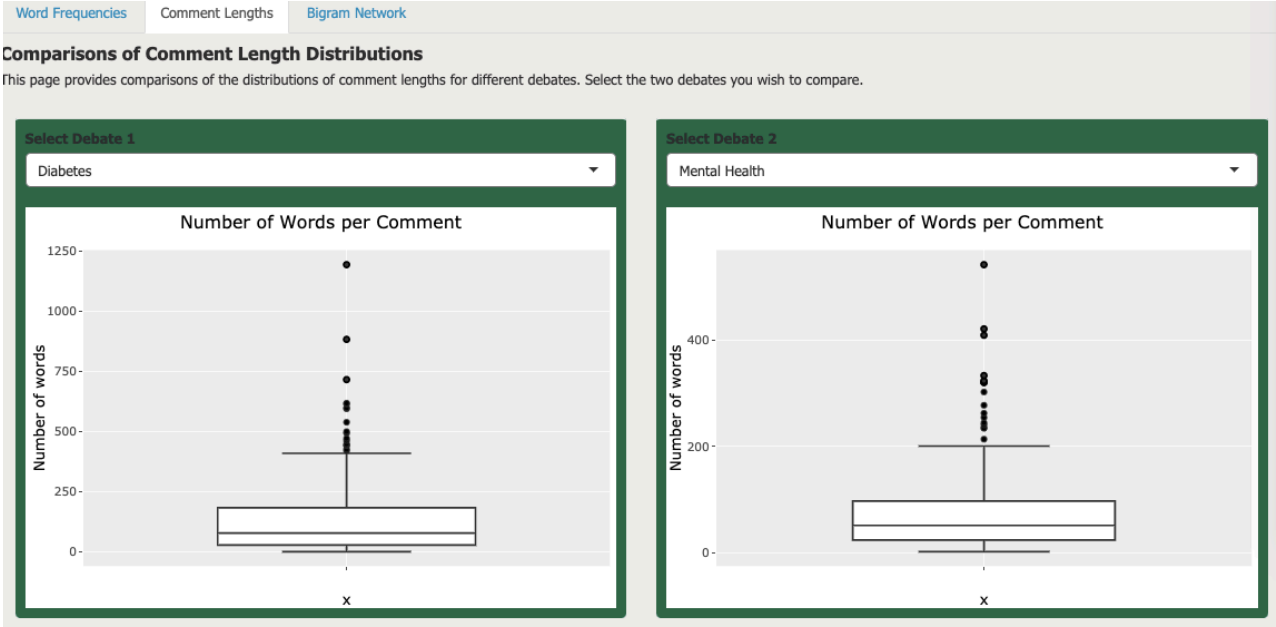
7.2.2 Comment Lengths

This subpage includes distributions of comment lengths for two different discussions. The user can select two files from the dropdown menu, and a boxplot of the number of words per comment in the discussion will be displayed below, side by side. The boxplots are interactive through the plotly function in R, which allows the user to hover over the plot to view the raw values for each quartile, whiskers and any outliers. This subpage was included in TheGist after consultations with members of the Digital Engagement team who wanted a way to view how substantial the comments being left in a discussion were, and an ability to track this over consecutive discussions. For example, this has been used to track differences in comment length from discussions over various topics, or discussions about the same topic at different times of the year. A Key Performance Indicator (KPI) for the team is to increase the substantialness of comments over a year or a parliamentary session[43], so including this subpage allows them to easily assess how much participants are writing in their comments.

Displaying the information in a boxplot was the preferred visualisation for the team and allows them to see more information than solely an average comment length. For example, although two discussions may have similar averages the distributions of comment length may differ heavily between the two, with one having a very small range and another having a larger range with more outliers. It may then be of interest to the team to isolate the longer comments and explore them further. However, his type of analysis must be used with caution when comparing discussion across platforms that have different limits on number of characters in a post, such as Twitter as this could result in misleading interpretations.

[43] Personal communication, Westminster, 2019

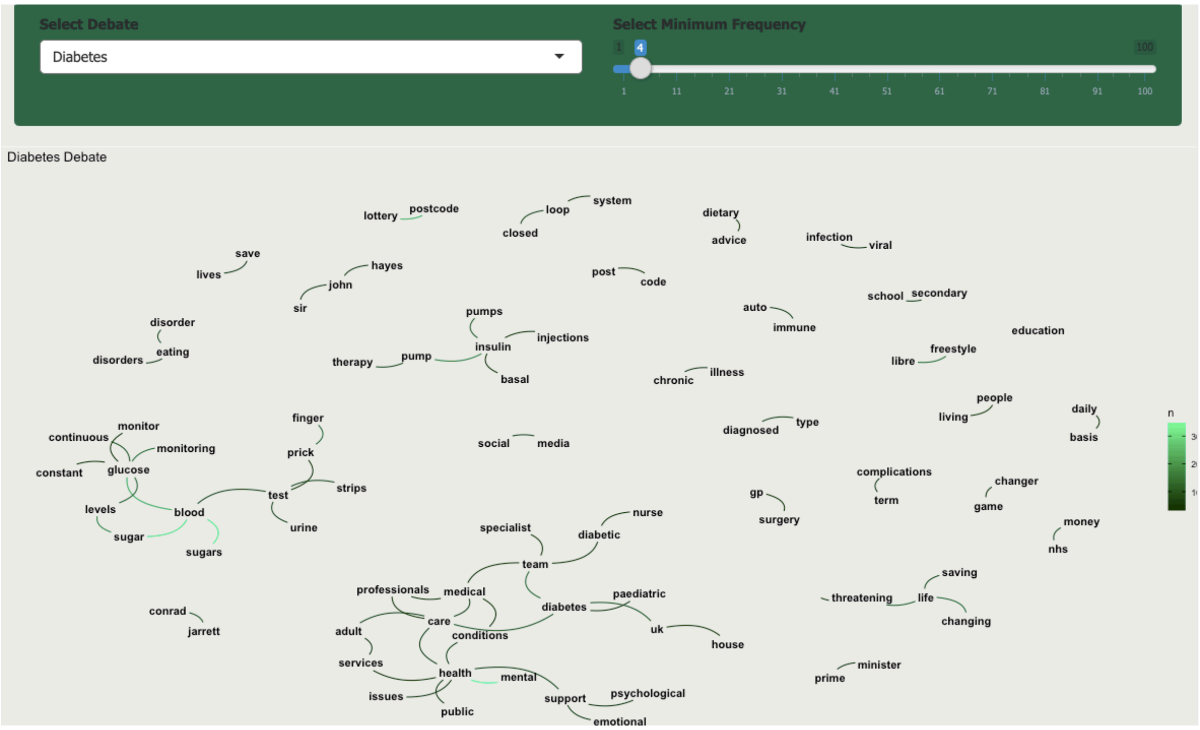
Figure 87: TheGist Comment Lengths snapshot



7.2.3 Bigram Network

The final subpage in this section provides a network of the most frequent consecutive word pairs (bigrams) in a discussion. As with the Word Frequencies subpage, there is a dropdown menu for selecting the discussion file and a slider bar for selecting the minimum frequency of bigrams to be displayed in the network plot (see Figure 88). The network is displayed with darker edges signifying a higher frequency of the bigram in the discussion. This type of analysis and visualisation was used throughout this project as a method of summarising the most

Figure 88: TheGist Bigram Network snapshot



popular themes in a text based on how often the words have been used together. This gives the user an insight into the co-occurring words which are most prominent in the debate in an intuitive and easily interpretable way. The advantage over simple keyword frequencies is that the relation between words creates context for keywords and allows for a better understanding of underlying themes. The decision moreover to visualise the bigrams in a network allows a user to easily view the most frequent bigrams while also seeing how various bigrams are linked to each other or even to build thematic cluster, something that would not be possible in a traditional bar plot.

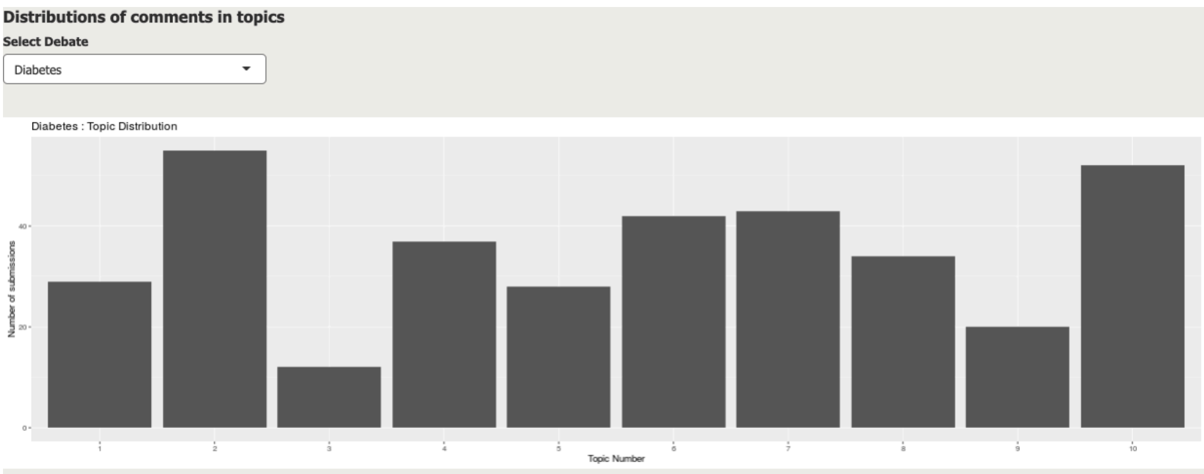
7.3 Topic Extraction

The second subpage in this section analyses the topics identified by Latent Dirichlet Allocation (LDA). Topic modes can be notoriously difficult to interpret due to the many options in creating the model (Knispelis, 2016). For example, the unsupervised method of LDA requires the analyst to pre-select the number of topics in the model and the chosen number of topics can greatly alter the method's effectiveness and the interpretation of the topics. There are different methods for optimising this value of k (number of topics) including those mentioned in section 3.2.3, however the ultimate decision lies in how easy the topics are to interpret based on the collection of words found in each. For example, some measures may find 20 topics to be the optimal value, but closer inspection of the words may show that several topics are redundant or that 20 topics is unmanageable and unhelpful for the user. Throughout the project and in the application, the LDAtuning package from R has been used to determine the optimal number of topics in a discussion.

7.3.1 Comments Distribution Across Topics

This subpage includes a dropdown menu of imported files and a bar plot of the distribution of comments across all the identified topics (Figure 89). This allows the user to explore whether the comments in a discussion are categorised into many different topics by LDA or only a selected few topics and to identify the most prominent topics in a debate

Figure 89: TheGist Topic Distribution snapshot



7.3.2 Topic by Words

The use of two visualisations (LDA and LDA-IDF) also allows for different interpretations of the topics to be analysed by the user to gain further understanding of the discussion. LDA has many advantages and is used widely in natural language processing, however a disadvantage is its lower performance on short text documents. Nevertheless, as this application can be used for both long and short text documents, Figure 90 shows the words distribution for each topic with words ranked by their highest beta scores as is the default for LDA.

The second visualisation on this subpage is the LDA-IDF plot. This uses the LDA topic model, but plots the words with the highest tdf-idf scores in each topic rather than the words with the highest beta scores (Figure 91). Occasionally, very frequent words in a discussion will also have the highest beta scores within each topic in the model, making interpretation of these topics more difficult for the user. The tf-idf measure gives precedence to words which are most unique within each topic, leading to an easier interpretation of the words which are in the topics but are not in other topics.

Figure 90: TheGist Topic by words (LDA) snapshot

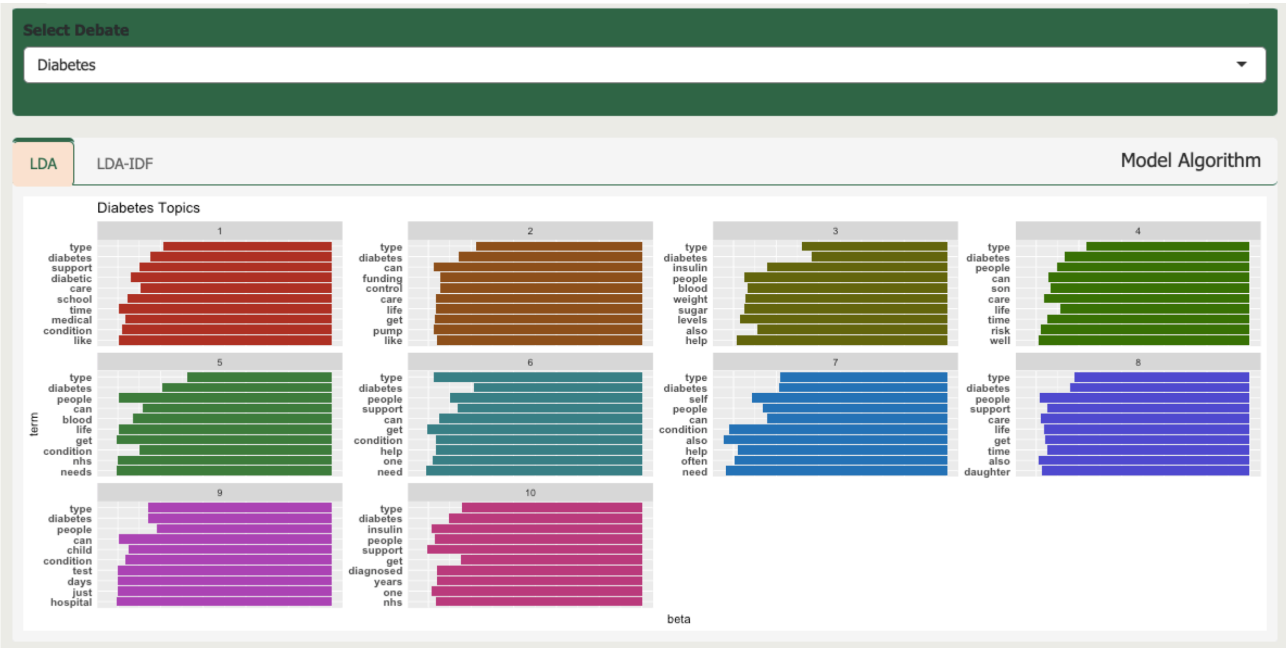
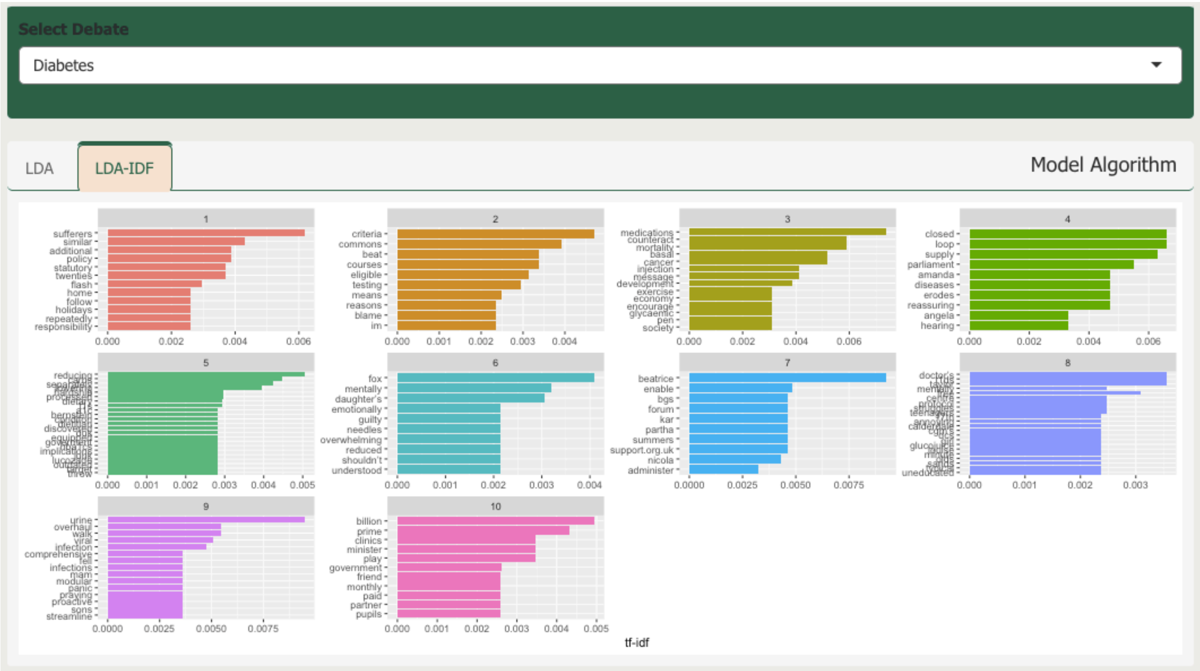


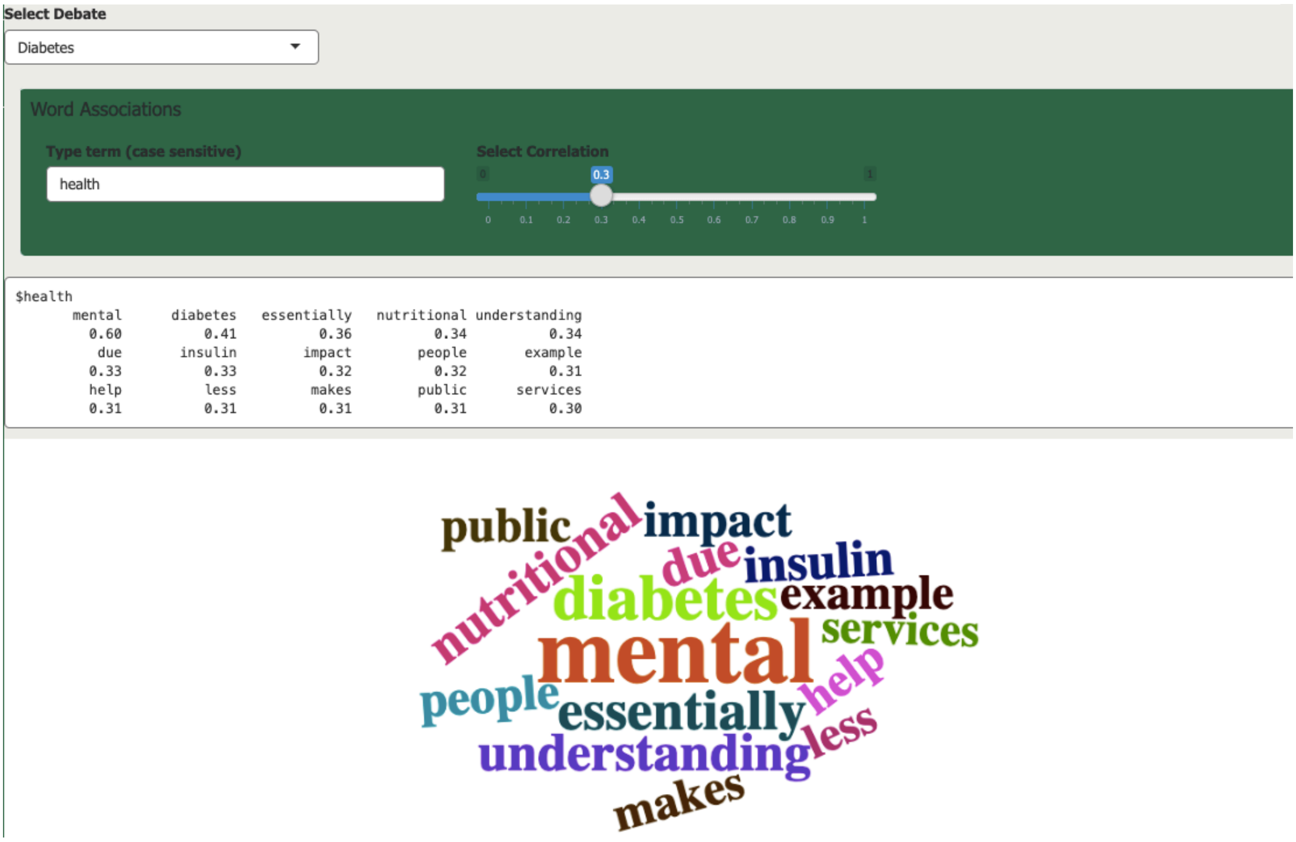
Figure 91: TheGist Topic by words (LDA-IDF) snapshot



7.3.3 Topic Associations

Staying with the theme of topic extraction, this subpage provides a word cloud of words which are most associated with each other. In the previous subpage, a user could select a discussion and view 10 words within that topic, however this subpage allows the user to choose any word within a discussion and view all other words which are used alongside it based on the DTM. There is a slider bar to select the correlation (from 0.0 to 1.0) indicating the lower limit. Words with at least a minimum correlation set by the slider are displayed in a text box and below in a word cloud, where words with a higher correlation to the search term are largest. For example, in a Facebook digital discussion about diabetes, when searching the term ‘health’, the words with the highest correlation are ‘mental’, ‘diabetes’, and ‘essential’, while the search term ‘sugar’ shows ‘alternatives’, ‘blood’, and ‘restrict’ (Figure 92). This feature is used to explore how participants speak about specific issues within a discussion and allows the Digital Engagement team to gain a deeper understanding of the primary concerns of the public in an easy way.

Figure 92: TheGist Topic Associations snapshot

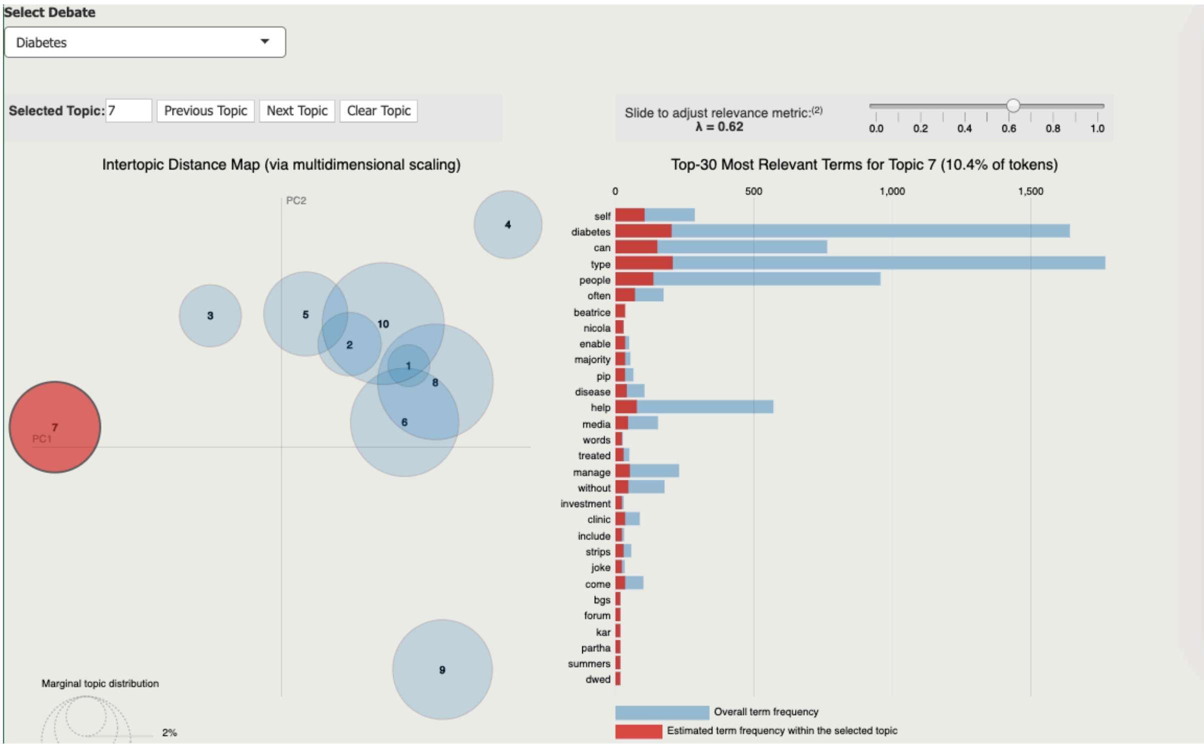


7.3.4 Topic Visualisations

To make the most out of the LDA model, the LDAvis package provides a clear visualisation of topic models which is easy to interpret (Sievert and Shirley, 2014). This visualisation is used to display the distances between topic in terms of the words they share, and a plot of the common words in each topic. It consists of two sections: (1) an inter-topic distance map showing topic distances from each other based on the types of words in each. This section uses principal component analysis to create a separation of topics across two dimensions, with topics close to each other signifying similarity in content; (2) a bar chart showing the top 30 words in each topic and overall in the dataset for each discussion. This is described in more detail in section 3.2.3.

As with the majority of the visualisations in TheGist, this is an interactive plot where the user can hover over each topic in the left-hand distance map and the 30 most relevant terms in the topic are displayed in the right-hand bar plot. A user can also select a word in the bar plot which will highlight the topics which that word also appears in on the left-hand side. This is a useful feature for exploring how topics are alike and which words they share. Where the Topic Distributions subpage just displays how many documents are in each topic, the LDAvis visualisation also shows if and how the topics interact with each other (Figure 93).

Figure 93: TheGist Topic Visualisations snapshot



7.4 Sentiment Analysis

The final section of TheGist analyses the sentiments of a discussion across three different sentiment lexicons discussed earlier. These are the Bing lexicon of positive and negative words, the Afinn lexicon of sentiment scales ranging from very negative -5 to very positive +5, and the NRC lexicon containing 8 sentiment categories. These three lexicons were used throughout the project to explore how participants express themselves online, in sections 6.4 and 6.6.2. Three lexicons are used as they provide slightly different interpretations of the text, which the Digital Engagement team use for different scenarios. The Bing lexicon is often used for a quick understanding of the general word use of participants and to see whether the general discussion is more positive or negative and by how much. For a closer look into the discussion, the NRC and Afinn sentiment lexicons are used to provide a detailed account of an engagement activity. The NRC lexicon contains sentiment and emotion categories which may not be helpful in certain scenarios, in which case the Afinn lexicon is used. For example, discussions which tend to attract very emotive comments such as those relating to animals are often analysed through the NRC lexicon which can isolate words belonging to particular emotions such as anger or joy. On the other hand, discussions which are more fact-based may not benefit from such a specific categorisation and instead use a numerical scale of sentiment polarity such as Afinn. Throughout this section, the user can switch between the three sentiment lexicons to gain a more thorough understanding of the discussion.

7.4.1 Distribution

This subpage displays a barplot of the percentage of comments/submissions in each sentiment category for each sentiment lexicon (Figure 94, Figure 95 and Figure 96). This feature of sentiment distribution is useful for understanding which sentiments are dominant in a discussion and can also be compared across the three sentiment lexicons to evaluate how they compare to each other.

Figure 94: TheGist Sentiment Distribution Bing snapshot

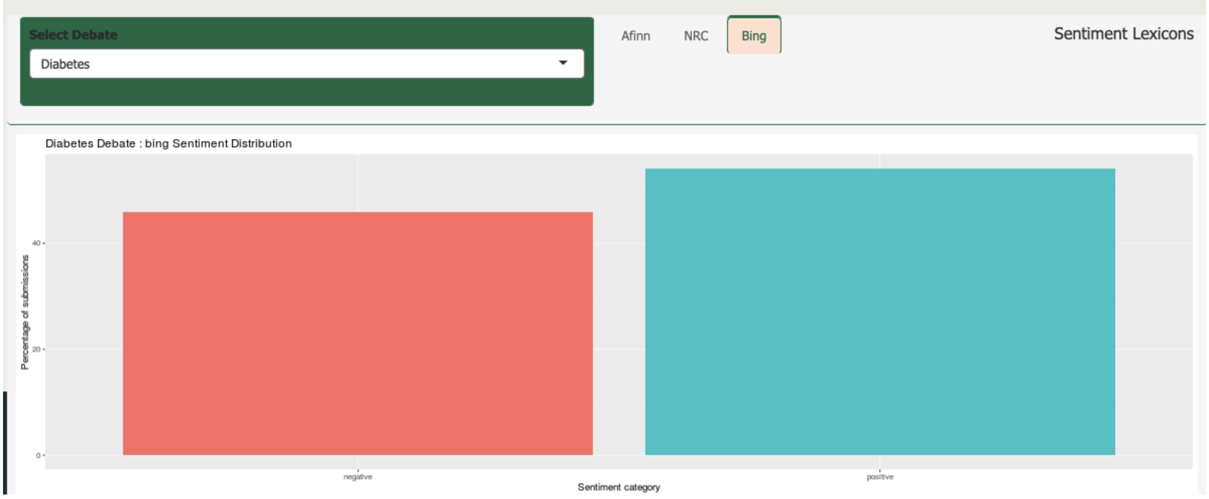


Figure 95: TheGist Sentiment Distribution NRC snapshot

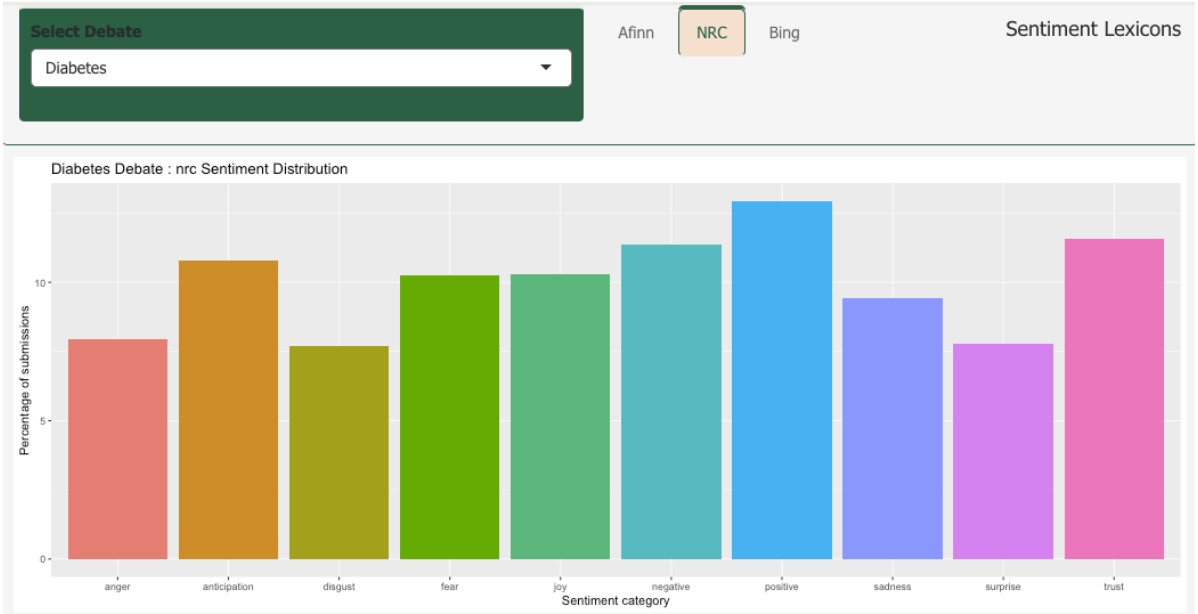
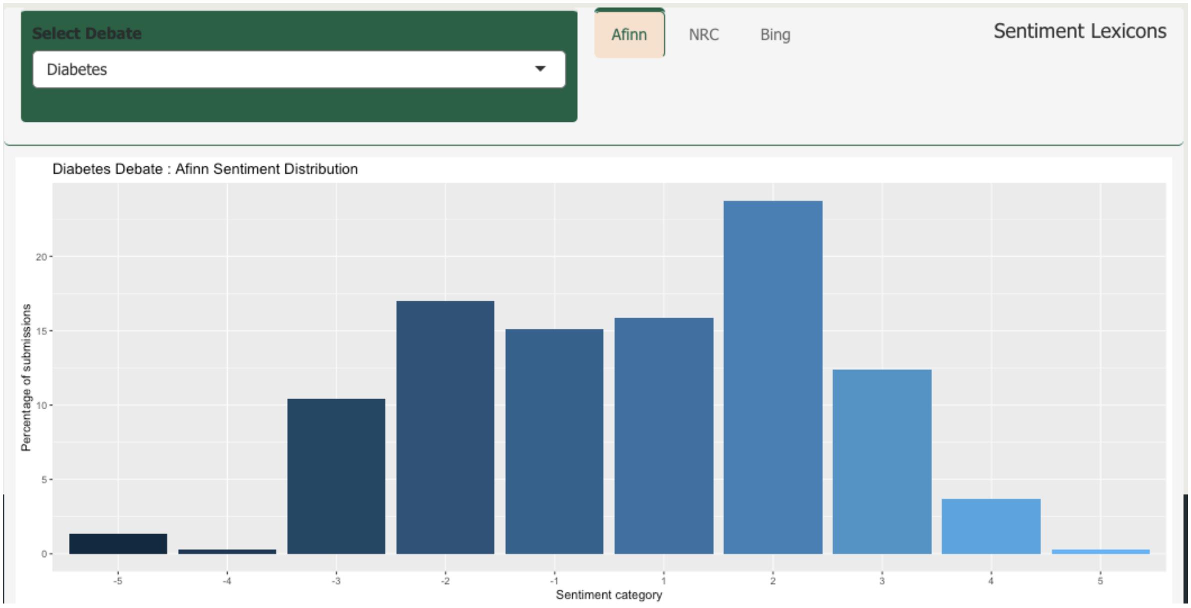


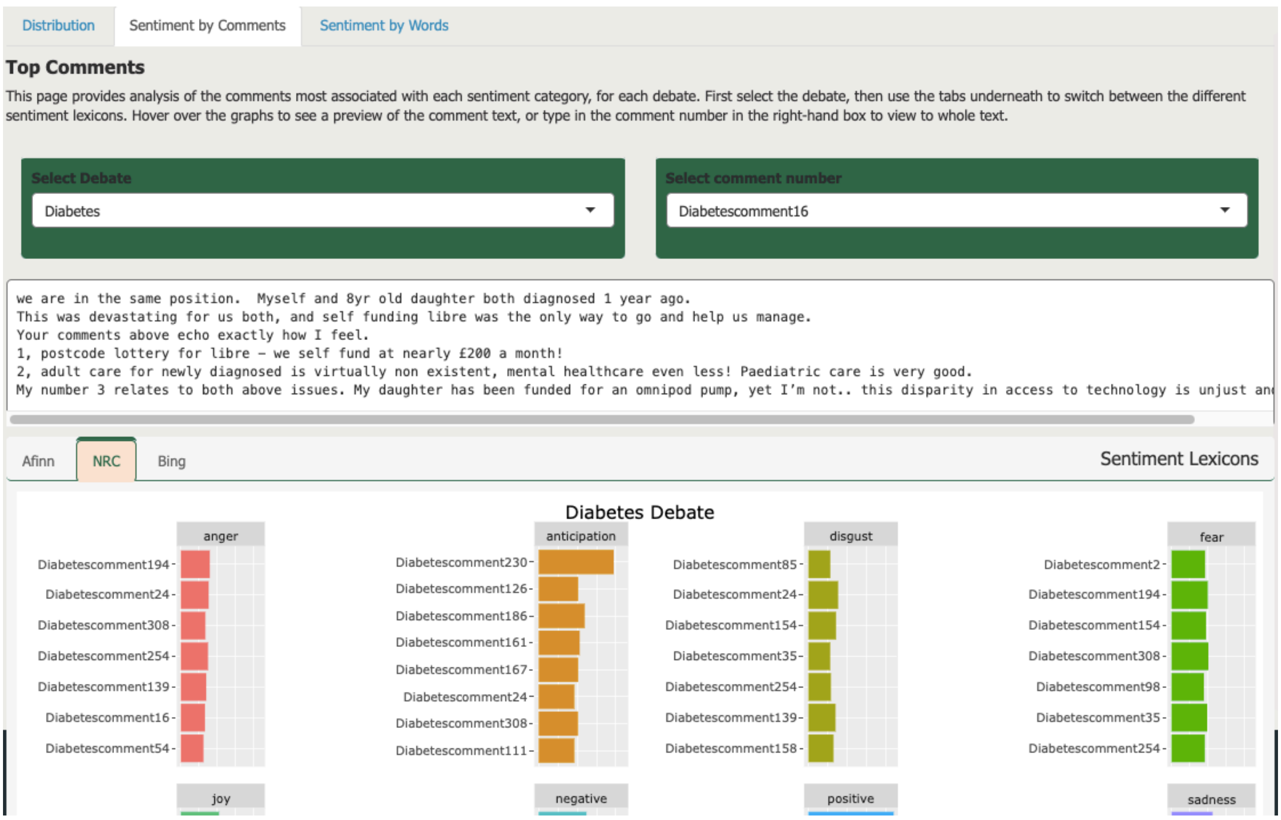
Figure 96: TheGist Sentiment Distribution Afinn snapshot



7.4.2 Representative Comments by Sentiment

This subpage contains barplots of the comments which have the highest contribution to each sentiment category (Figure 97). This is an interactive plot where the user can hover over each

Figure 97: TheGist Comments by sentiment



bar revealing the comment number, the percentage of words in the comment belonging to that sentiment category, and the comment text. There is also a combined search and dropdown menu where the user can look for a specific comment number within the discussion, and the full comment text will be displayed in a box underneath. This feature is added to TheGist because some comments are very long as identified by the Comment Lengths subpage in the General section and the Digital Engagement team may wish to look further into this.

7.4.3 Representative Words by Sentiment

The third subpage of the Sentiment Analysis application section of TheGist gives plots of the words appearing in each sentiment category. The Bing lexicon with only two categories of positive and negative is displayed in a word cloud with the larger words contributing more heavily to the colour-coded sentiment category. The red text represents negative words while the green text represents positive words (Figure 98). For Afinn (Figure 99) and NRC (Figure 100) lexicons which have many sentiment categories this information is displayed in barplots.

Figure 98: TheGist Sentiment by words Bing wordcloud

Figure 99: TheGist Afinn Sentiment by words

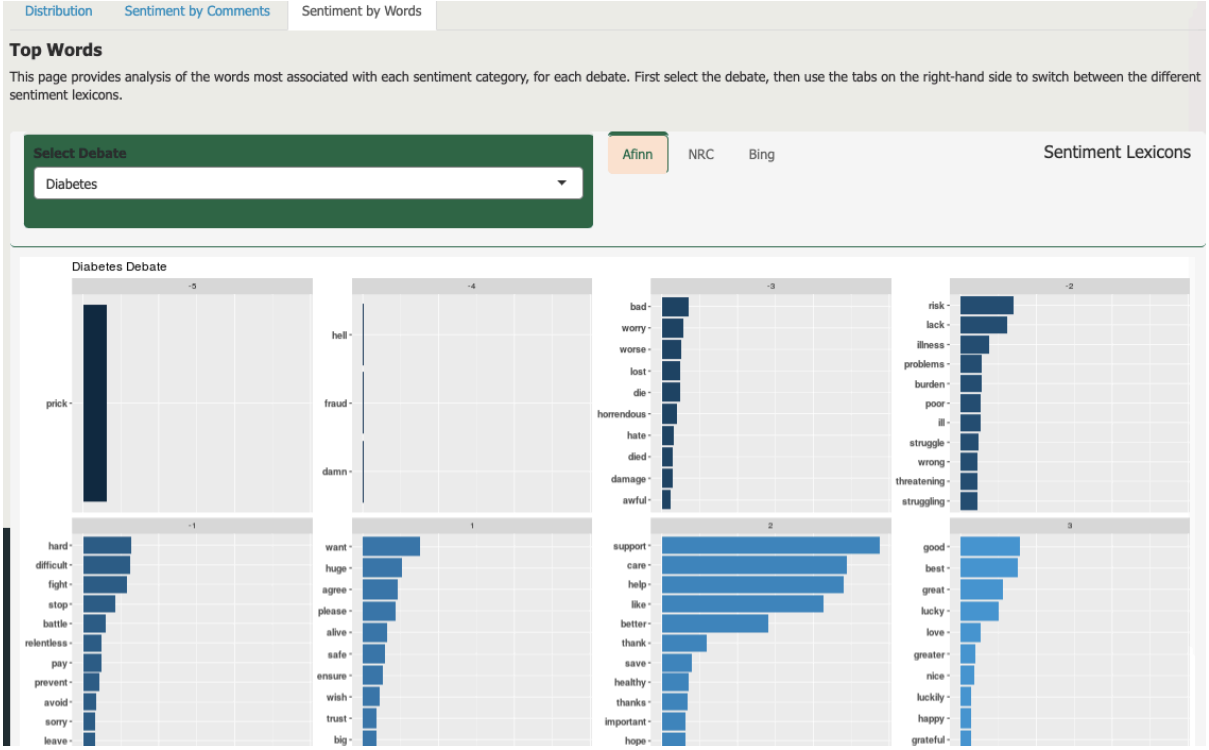
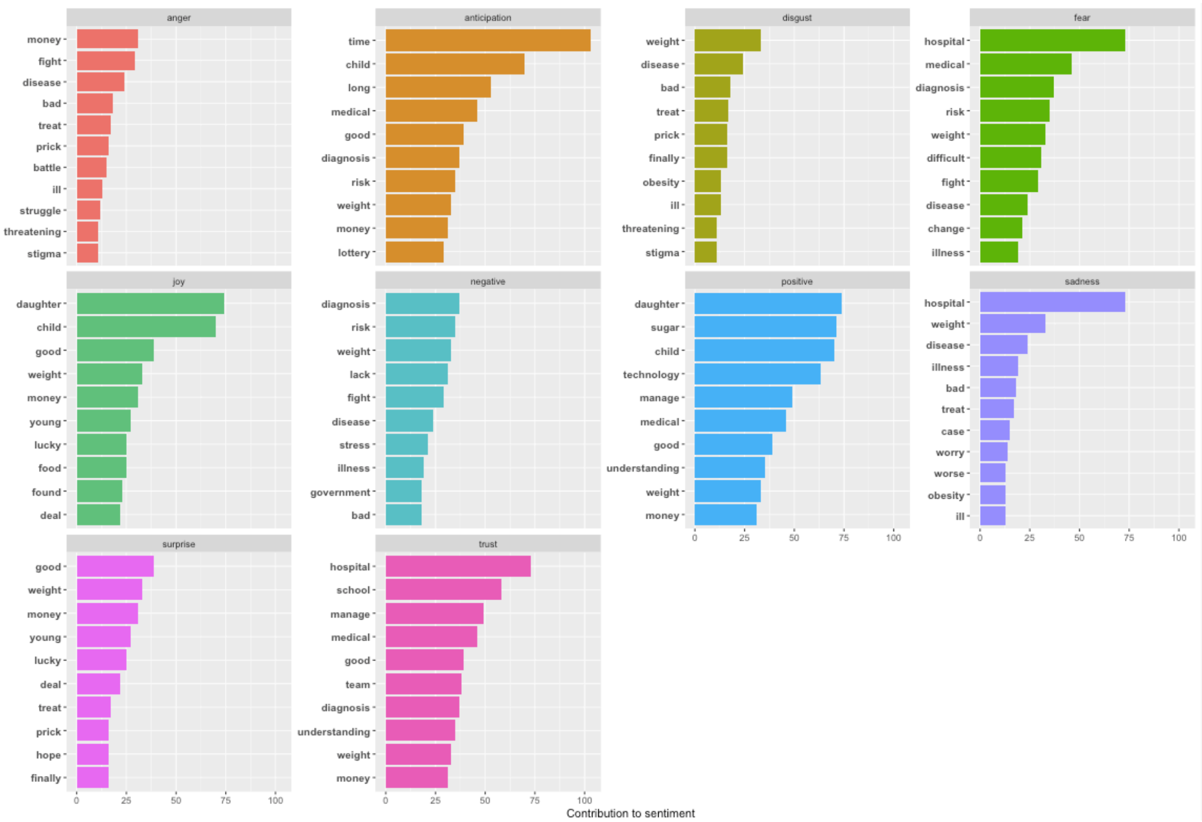


Figure 100: TheGist NRC Sentiment by words



7.5 Implementation of the application

As described in section 4.3, once the application was created there was some difficulty implementing it into Parliament so the Digital Engagement team were able to use it. The application was built through Shiny in R studio which has a feature to deploy applications to the cloud, allowing users to access the application through a webpage. However, from speaking with the Digital Engagement team, it became apparent they were keen to have more exclusivity of the application including it being password protected. This is possible through the Shiny cloud however comes at a monthly cost. The Digital Engagement team were unable to allocate the necessary funds from their budget for the application[44] and there is a lengthy accreditation process for new software in Parliament, so it is currently only accessible to them through the free version. While this does not raise too many problems at present, it does restrict the team from uploading comments from private or sensitive discussions, or closed surveys to ensure they are not accidentally made public. The use of the free version of the shiny cloud also restricts the number of users on the application at one time meaning only one member of the team is able to use the application at once.

The difficulties in budgeting and speaking to the relevant people and departments to make the application usable were further exacerbated by the current Covid-19 pandemic, which meant that this use could not been further progressed. There is hence a delay in the implementation of TheGist for the Digital Engagement team. However, I am still in contact with the House of Commons and currently working to get the application accredited and available to the teams who it was created for.

[44] Personal communication, 2019, Westminster

Chapter 8 Discussion and Conclusions

The main focus of this thesis is to understand how digital engagement with the UK Parliament can be evaluated and harnessed in the most effective manner. This chapter concludes this thesis, highlights the main findings with respect to how digital engagement is conducted and analysed, and reflects on how the institution can enhance engagement activities. The following sections provide a summary of the answers to the research questions (examined in section 1.1.2) and a more general reflection of the project outcomes and outlooks.

8.1 Summary of findings

In order to measure the effectiveness of public engagement activities, I explored how public engagement is categorised between activities which inform the public of parliamentary business and activities which encourage the public to actively participate in proceedings. While there are many different interpretations of engagement, through analysis of the existing literature on parliamentary engagement I find a way to conceptualise engagement into its different components. I use a framework which incorporates the flow of engagement and whether the public has an active role in the activity. This divides engagement into two branches: left, where the flow of information is one-directional out of the institution, and right, where the flow alternates between the public and the institution. This conceptualisation closely mirrors the current practice of engagement in the UK Parliament and how parliamentary services are organised.

However, these dimensions of engagement are measured differently depending on who is responsible for the activity. Chapter 4 explored the organisational structure of various teams responsible for digital engagement in the UK Parliament and found that there are many teams doing different types of engagement and for different purposes. Internal organisational structures of departments and teams influence how engagement activities are conducted and how their success is measured. Select committees in particular have a unique consensual approach, placing them in a position within the institution allowing them to tackle a range of different issues and explore different methods of engaging with the public. This flexibility of select committees became very useful when it came to conducting demonstration tests on digital engagement platforms later in the project, and highlights the different abilities of various teams in Parliament when organising engagement activities.

Once I understood the context of engagement and who was involved, I put the methods of evaluation into practice using exclusive data from the various UK Parliament social media accounts (Chapter 5). These channels are used primarily for the left branch of the spectrum where information is flowing from the institution to the public and the primary aim is to inform the public of parliamentary proceedings. I found a range of users from different parts of the UK engaging with Parliament using Twitter and identified cities where users were the most active and concentrated. I also identified specific House of Lords select committees whose Twitter followers suggested a high number of bots and explored the inequalities in proportion of users following each committee account. This suggests an uneven method of managing these accounts causing potential confusion among the users. A closer evaluation of the UK House of Commons' posts on Facebook shows how the public are enthusiastic in responding when asked for their opinions in the form of Facebook Digital Discussion or Debates but the type of post made has an effect on this engagement. Links to external websites are not engaged with as much as photos or embedded videos. This suggests the user-journey is an important component of digital engagement and can create barriers to the public engaging. A simpler user journey where users were not directed to another website in order to access information (such as with photos) elicited higher levels of participation from the public, and this was incorporated into my demonstration tests with Discourse. The existing evaluation of the most popular posts in terms of number of comments or likes failed to show the topics which encouraged the most engaged users (defined by those who repeatedly commented on a particular post or topic). Through my evaluation measure, I also found differences in the subject matter of posts which the public interacted with. By expanding this evaluation using topic modelling, I showed how users interacted with the different topics posted by the House of Commons and revealed issues which are important to the public.

Addressing the activities initiated and controlled by Parliament to inform and educate the public evaluates the performance of only one type of engagement. We need to expand this evaluation to understand how the public responds to these acts of engagement by the institution and what Parliament can learn from their citizens. An examination of several digital debates held by the UK Parliament showed how computational social science methods were applied to digital engagement data to provide another element of evaluation. Specifically, I used topic modelling to algorithmically extract the main themes from large volumes of textual data to understand participants' views on certain issues. This demonstrated the many different sub-topics involved in a single discussion along with which of these sub-topics participants were most passionate about by the number of comments categorised into each sub-topic. K-means clustering based on socio-demographic identifiers in text also provided insight into how participants from different backgrounds approach an issue. This is helpful for my collaborators in the UK Parliament who wish to understand if their engagement activities are reaching a varied range of citizens and learn how different age groups react to the same issues. This experimental approach to estimating participants' backgrounds provides a method of evaluating the success of an engagement activity based on the types of people involved and helps the teams responsible for engagement to reflect on their digital work.

Parliament's digital engagement activities are effective at bringing groups of people together based on the subject matter of posts they interact with. Social network analysis of user activity with Facebook posts showed that similar topics are the glue which hold the online community together. Likewise, when analysing select committee evidence networks, submitters (whether individuals or organisations) are more alike in the topic of inquiries they post rather than in the existing committee structure. These effectiveness measures focus on who is participating, however there have also been achievements with the journey of engagement. During the demonstration tests with the Discourse platform, I found users to transition from signing a petition (Consultation), to participating in an online discussion forum (Discussion) to submitting formal written evidence to the committee.

Along with identifying various internal barriers to engagement, I also explored how the existing channels for online engagement could hinder the engagement process. Moderation and the ability to segment discussions into smaller topics was an important factor in controlling the volume and interpretations of comments on the parliament's side. This was also important for the public's side in terms of directing them to the specific areas of the discussion which they were most interested in. To put these ideas into practice, I trialled a new engagement platform, Discourse, using live select committee inquiries. These demonstration tests showed a difference in how participants act on an existing engagement channel used by Parliament, such as Twitter, compared to Discourse based on the different features used in each. The separation of topics, ability for users to moderate themselves and the flexibility of data exports to track user interactions allows me to extract much more useful information from the platform. I found participants were more solutions-focussed on Discourse compared to Twitter when given a space to deliberate with each other and split the conversation into distinct parts. Users were also less extreme with the way they expressed themselves on Discourse where a range of sentiments were expressed. Conversely, Twitter had a higher proportion of tweets expressed by a single sentiment.

Specific logistical issues arose in one case at the end of the demonstration test. Each committee received a csv file of all comments posted to the Discourse discussion and I produced an interactive HTML report outlining the main themes and sentiments expressed across both platforms for their discussions. While for the most part, this format was very well received, in one case the committee asked for the results to be given in Word or PDF format so they could be submitted as formal evidence to the committee's inquiry. This was done but as a result lost the interactive feature of the HTML. More importantly, this suggested the committee did not fully recognise the purpose of producing reports in this way, namely to step away from the usual way of submitting evidence and encourage the use of different forms of analysis into Parliament. Nevertheless, the report was used in the final published inquiry report, however this is an interesting insight into the perceived acceptance of new technology and methods by committees. It suggests a potential lack of understanding or willingness from all committees to adapt in practice, alongside an absence of internal processes that would facilitate different formats of evidence to be submitted. This limitation of the types of evidence which can be submitted was raised in the Liaison Committee (2019) report on the effectiveness of select committees. They acknowledge that the current restrictions on format for evidence are out-dated and reference other parliaments who have expanded their processes to accept video and audio format. They therefore recommend that "whatever its medium or format, information submitted to committees which they then seek to publish by order of the House ought to be recognised as formal evidence. The House must take the necessary steps to bring about this change." (Liaison Committee, 2019, p. 45).

Following the Environmental Audit Committee (EAC) discussion about invasive species, 51 formal written submissions (16% of total participants) were made to the Invasive Species inquiry by participants who wanted to go into further detail and provide their experience and expertise[45]. This action in such a large quantity showed an escalation of engagement from petitions or social media, to an online discussion, to submitting written evidence. Such a pattern is something that many teams dealing with engagement are keen to see, but generally difficult to obtain. And it was due to this engagement process that the questioning of the Parliamentary Under Secretary of State for Rural Affairs and Biosecurity was altered to accommodate the concerns raised in the Invasive Species discussion about the welfare of squirrels. This is a direct indication of meaningfulness of the engagement session and recognition by officials, and shows a direct link between public input – especially one that could be described as argumentative – and policy making.

I also focus on combining what I have learned from a project that has both academic value and institutional value to the collaborators in the UK Parliament. Through conversations with parliamentary officials over the three years of the project, I found that data analysis was a major internal barrier to engagement and many teams stopped conducting online engagement sessions due to an inability to manage the large volume of comments. This fear directly contradicts many of the key performance indicators and measures of effectiveness of some teams regarding digital engagement which focus on encouraging a high number of participants and as a result, comments. Therefore, a key aim in understanding how citizen engagement can be incorporated into decision-making is about facilitating the analysis of textual data and reduce this internal barrier to engagement.

In the previous chapter, I introduced TheGist web application which brings many of the data analysis methods used in Chapter 5 and Chapter 6 together into an easy-to-use tool which does not require any prior programming skills. The features of this application are explained

[45] Personal communication, Westminster,

in detail in Chapter 7 and aim to give a condensed summary of an online discussion or survey using established natural language processing techniques and easily interpretable visualisations. This tool reduces the time needed to understand the main issues raised by participants in the discussion while also providing information on how substantial the comments were and the sentiments participants used to express themselves. This can therefore condense a discussion containing several thousand comments into visualisations which can be interpreted by the Digital Engagement team, as well as Members of Parliament who initiate the debates.

In summary, this research has provided an answer to the three questions outlined in section 1.1.2. I have explored how public engagement can be defined into its various dimensions based on who is initiating the activity and the flow of information, as well as how different areas of parliament can effectively evaluate their activities. I have explored the nature of citizen input from digital discussions on Facebook and Twitter to reveal the varied subject matter and the way participants express themselves online, especially when discussing something they are passionate about. Finally, I have designed demonstration tests to cover the life-cycle of an engagement activity, putting emphasis on how the citizen input can be meaningfully incorporated into policy making, by using purpose-built tools and ensuring clarity at every stage.

As a result of this research, several teams within Parliament have already begun to work more closely together on mutual goals, for example the Digital Engagement and PDS teams. During my PhD and regular presence in the UK Parliament, I have been trusted to conduct engagement experiments on live select committee inquiries and have been approached by other teams and committees wanting to improve their methods of analysis for digital engagement. I have also been asked to provide input to the UK-wide strategy for engagement within Parliament over the next five years.

8.2 Final reflections

The collaborative nature of this PhD gave me insights into the day-to-day routines of the teams responsible for managing Parliament's engagement with the public. This granted me exclusive access to data, and allowed me to build and maintain working relationships with different officials. This ensures my research is conducted with the stakeholder in mind and that any findings, conclusions, or recommendations are developed in consideration of both academic rigour and practical applicability in the workplace. This research has touched on many aspects of digital engagement within the UK Parliament and while it is clear there is no 'one size fits all', there are some technical and institutional factors which should be considered in any case of digital engagement.

Clarity at the beginning of an engagement activity to identify the primary aims and motivations behind conducting the activity are paramount to its success. This means the staff understand how to measure whether the activity has been successful or not, and the public are aware of how their input will be used and what level of involvement they can expect from the institution. Closing the feedback loop between Parliament and the public during engagement sessions will also reduce frustrations from the public at not feeling heard and encourage them to participate in the future.

Analysing existing digital engagement data and conducting demonstration tests with select committees allowed me to put these ideas of clarity in the aims and purpose of an engagement session and how it will be evaluated, into practice. In other words, truly understanding the reasoning for conducting engagement rather than doing engagement just for the sake of it. The reflections of this project can therefore be encompassed and demonstrated through a critical view of the demonstration tests. Throughout this project, potential limitations and elements to improve in the future became apparent, and can be categorised as those related to the technical aspects of the digital platforms used, and those arising due to internal processes within Parliament. It is vital to consider both the technical and institutional factors when evaluating the effectiveness of public engagement, as the value of these demonstration tests and in turn, the value of digital engagement, is dependent on several factors being thoughtfully considered and successfully working together.

However, what has become clear is that addressing these internal institutional barriers first is vital to successful public engagement strategy. These range from inter-generational challenges among officials, to the many procedural processes, to issues arising due to the unique political climate the UK found itself due to Brexit. On the 16th June 2016, the public voted to leave the European Union in a nationwide referendum. Although occurring years prior to my project, Brexit has continued to have a strong effect on parliamentary business and made conducting engagement and demonstration tests more difficult. For example, May-June 2019 was a key period in the Brexit negotiations and featured Prime minister Theresa May's resignation (BBC News, 2019). The committee inquiries involved in the demonstration tests were not overly affected by the political climate, meaning the committees could set their own agenda and still continue accepting evidence for their inquiries and the experiments could go ahead as planned. Whereas the select committees were not too affected by Brexit, the same cannot be said for other areas of Parliament. This research is collaborative with the House of Commons, but specifically the Digital Engagement team who sit under the Participation department. The data leading to analysis of digital debates on the UK House of Commons official Facebook page detailed in Chapter 6 is managed by this team, and the demonstration tests were originally intended to be carried out also through them. However, their digital debates are much more dependent on the MPs deciding to have an online discussion based on the business part of Parliament. This means their work is more heavily affected by the uncertainty that Brexit has caused to the parliamentary business schedule. Firm plans and schedules of inquiries and discussions were required to prepare the different components of the demonstration tests, and therefore working with the select committees proved a more feasible choice. This highlights the importance of understanding where each team with a remit for digital engagement sits within the wider scope of parliamentary processes.

Nevertheless, I successfully conducted demonstration tests analysing direct comparisons between platforms in live select committee inquiries and these were included in the final published inquiry reports (Environment Food and Rural Affairs Committee, 2019; Transport Committee, 2019). The novel use of Discourse as a discussion platform within Parliament was also mentioned in the Liaison Committee's inquiry report on the effectiveness of select committees (Liaison Committee, 2019), showing that although trialled in just three inquiries, the use of new technologies and approaches is being recognized across Parliament as something to be considered when communicating with the public.

While I can explore different ways of mitigating the negative effects of digital engagement such as trolls and argumentative users, care must be taken not to unintentionally detract from allowing participants to naturally engage with each other. It is a normal human behaviour to disagree and argue with each other, so conducting engagement sessions online must find ways to allow people to be their normal selves while ensuring any disagreements are not so negative or offensive that they derail the conversation as a whole. This is achieved through using tools and platforms specifically designed to encourage deliberative engagement while allowing participants control over their conversations. Combining the disciplines of data science and social science allows us to use a range of quantitative methods specific to the data types (in this case text and social media) used by my collaborators in Parliament. And creating an application such as TheGist makes sure the methods to evaluate engagement and reduce internal barriers are sustainable beyond the lifetime of this PhD.

In summary, this research shows that while we can design an online engagement activity and create effective methods to evaluate the success of the activity, there are still institutional matters to consider which can cause difficulties in any activity. Addressing external barriers to engagement which directly affect the public such as ease of use of the platform and different features which enable segmentation of discussions are important factors in engagement sessions. However, addressing the internal barriers which directly affect how the institution views, organises, and incorporates digital engagement into their daily practices is crucial to the effectiveness and longevity of parliamentary digital engagement. These barriers are the internal processes of how the public's views can be incorporated into decision -making, if at all, what exactly officials are expecting to gain from reaching out to the public, and crucially how they plan to evaluate engagement to capture as much of the public's voice as possible. Without this understanding and clarification from the outset, digital engagement activities will continue to address only part of the problem, mostly focussed on the public's access to participatory activities, rather than the institution's readiness to manage them.

References

Adeva, J. G., Atxa, J. P., Carrillo, M. U. and Zengotitabengoa, E. A. (2014) 'Automatic text classification to support systematic reviews in medicine', *Expert Systems with Applications,* 41(4), pp. 1498-1508.

Aggarwal, C. C. and Zhai, C. (2012) *Mining text data.* Springer Science & Business Media.

AnimalAid (2019) *Speakup for squirrels!* Available at: https://www.animalaid.org.uk/the-issues/our-campaigns/wildlife/squirrels-2/speak-up-for-squirrels/.

Ares, E., Library, H.o.C. (2018) *Petition to ban the sale of fur in the UK*.

Arnstein, S. R. (1969) 'A ladder of citizen participation', *Journal of the American Institute of planners,* 35(4), pp. 216-224.

Arun, R., Suresh, V., Veni Madhavan, C. E. and Narasimha Murthy, M. N. 'On Finding the Natural Number of Topics with Latent Dirichlet Allocation: Some Observations'. *Advances in Knowledge Discovery and Data Mining*, Berlin, Heidelberg, 2010//: Springer Berlin Heidelberg, 391-402.

Asher, M., Leston-Bandeira, C. and Spaiser, V. (2017) 'Assessing the effectiveness of e-petitioning through Twitter conversations'.

Bartlett, J. (2018) *The People Vs Tech: How the internet is killing democracy (and how we save it).* Penguin.

BBC News (2018) *Italy PM-designate Conte gives up bid for government*. bbc.co.uk. Available at: http://www.bbc.co.uk/news/world-europe-44275010 (Accessed: 28 May 2018.

BBC News (2019) 'Theresa May quits: UK set for new PM by end of July'. Available at: https://www.bbc.co.uk/news/uk-politics-48395905.

Becker, S., Kurbjuweit, D., Muller, P., Rosenbach, M. and Theile, M. (2012) 'German Pirate Party Attempts to Reinvent Politics', *Der Speigel*.

Bedi, P. and Sharma, C. (2016) 'Community detection in social networks', *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery,* 6(3), pp. 115-135.

Belk, R. (2016) 'Extended self and the digital world', *Current Opinion in Psychology,* 10, pp. 50-54.

Blei, D. M. 2017. Prof. David Blei - Probabilistic Topic Models and User Behavior. *In:* The School of Informatics at the University of Edinburgh (ed.). YouTube.

Blei, D. M., Ng, A. Y. and Jordan, M. I. (2003) 'Latent dirichlet allocation', *Journal of machine Learning research,* 3(Jan), pp. 993-1022.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R. and Lefebvre, E. (2008) 'Fast unfolding of communities in large networks', *Journal of statistical mechanics: theory and experiment,* 2008(10), pp. P10008.

Blum, C. and Zuber, C. I. (2016) 'Liquid Democracy: Potentials, Problems, and Perspectives', *Journal of Political Philosophy,* 24(2), pp. 162-182.

Bonsón, E., Royo, S. and Ratkai, M. (2017) 'Facebook practices in Western European municipalities: An empirical analysis of activity and citizens' engagement', *Administration & Society,* 49(3), pp. 320-347.

Brandt, J., Buckingham, K., Buntain, C., Anderson, W., Ray, S., Pool, J.-R. and Ferrari, N. (2020) 'Identifying social media user demographics and topic diversity with computational social science: a case study of a major international policy forum', *Journal of Computational Social Science*.

Brundidge, J. (2010) 'Encountering "Difference" in the Contemporary Public Sphere: The Contribution of the Internet to the Heterogeneity of Political Discussion Networks', *Journal of Communication,* 60(4), pp. 680-700.

Burgess, J. and Bruns, A. (2012) 'Twitter archives and the challenges of" Big Social Data" for media and communication research', *M/C Journal,* 15(5).

Cao, J., Xia, T., Li, J., Zhang, Y. and Tang, S. (2009) 'A density-based method for adaptive LDA model selection', *Neurocomputing,* 72(7), pp. 1775-1781.

Carpini, M. X. D., Cook, F. L. and Jacobs, L. R. (2004) 'Public deliberation, discursive participation, and citizen engagement: A review of the empirical literature', *Annu. Rev. Polit. Sci.,* 7, pp. 315-344.

Cha, M., Haddadi, H., Benevenuto, F. and Gummadi, K. (2010) 'Measuring User Influence in Twitter: The Million Follower Fallacy', *Proceedings of the International AAAI Conference on Web and Social Media,* 4(1).

Chang, J., Boyd-Graber, J., Wang, C., Gerrish, S. and Blei, D. M. 'Reading tea leaves: How humans interpret topic models'. 2009: Citeseer, 288-296.

Charrad, M., Ghazzali, N., Boiteau, V. and Niknafs, A. (2012) 'NbClust Package: finding the relevant number of clusters in a dataset', *J. Stat. Softw*.

Chavkin, L. (1997) 'Readability and reading ease revisited: State-adopted science textbooks', *The Clearing House: A Journal of Educational Strategies, Issues and Ideas,* 70(3), pp. 151-154.

Chu, Z., Gianvecchio, S., Wang, H. and Jajodia, S. (2012) 'Detecting automation of twitter accounts: Are you a human, bot, or cyborg?', *IEEE Transactions on Dependable and Secure Computing,* 9(6), pp. 811-824.

Cihon, P., Yasseri, T., Hale, S. and Margetts, H. (2016) 'TWEETING FOR THE CAUSE: Network analysis of UK e-petition sharing activity'.

Citraro, S. and Rossetti, G. 'Eva: Attribute-Aware Network Segmentation'. *International Conference on Complex Networks and Their Applications*: Springer, 141-151.

Coleman, S. (2004) 'Connecting Parliament to the Public via the Internet', *Information, Communication & Society,* 7(1), pp. 1-22.

Coleman, S. and Firmstone, J. (2014) 'Contested meanings of public engagement: exploring discourse and practice within a British city council', *Media, Culture & Society,* 36(6), pp. 826-844.

Coleman, S. and Gotze, J. (2001) *Bowling together: Online public engagement in policy deliberation.* Hansard Society London.

Colleoni, E., Rozza, A. and Arvidsson, A. (2014) 'Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data', *Journal of Communication,* 64(2), pp. 317-332.

Commission, D. D. (2015) *Open Up! Report of the Speaker's Commission on Digital Democracy*.

Committee, L. (2015) *Building public engagement: Options for developing select committee outreach*, London: House of Commons. Available at: https://publications.parliament.uk/pa/cm201516/cmselect/cmliaisn/470/47002.htm.

Conroy, M., Feezell, J. T. and Guerrero, M. (2012) 'Facebook and political engagement: A study of online political group membership and offline political engagement', *Computers in Human Behavior,* 28(5), pp. 1535-1546.

Countryfile (2019) *Red squirrel guide*. Available at: https://www.countryfile.com/wildlife/mammals/red-squirrel-guide-how-to-identify-and-where-to-see-in-britain/ (Accessed: 30 September 2019.

Cozma, R. (2013) 'John H. Parmelee and Shannon L. Bichard. Politics and the Twitter Revolution: How Tweets Influence the Relationship Between Political Leaders and the Public', *Mass Communication and Society,* 16, pp. 460-463.

Dalton, R. J. (2004) *Democratic Challenges, Democratic Choices: The Erosion of Political Support in Advanced Industrial Democracies.*

Dean, N. 'Social Network Analysis'. *Advanced Quantitave Methods Taster Day*, University of Sheffield, 23 January 2018.
Digital Democracy Commission (2015) *Open Up! Report of the Speaker's Commission on Digital Democracy*.
Discourse.org (2019) *Civilized discussion for your community*. Available at: https://www.discourse.org/ (Accessed: 16 November 2019.
Dorey, P. and Purvis, M. (2018) 'Representation in the Lords', *Exploring Parliament*: Oxford University Press.
Dreiger, P. (2013) 'Semantic Network Analysis as a method for visual text analysis', *Procedia - Social and Behavioural Sciences,* 79, pp. 4-17.
Education Committee 2011. Parliament wants your view. The Student Room.
Education Committee (2012) *Michael Gove answers #AskGove twitter questions*. Available at: https://www.parliament.uk/education-committee-askgove-twitter-questions.
Edwards, A., Housley, W., Williams, M., Sloan, L. and Williams, M. (2013) 'Digital social research, social media and the sociological imagination: Surrogacy, augmentation and re-orientation', *International Journal of Social Research Methodology,* 16.
Empirica (2014) *e-Skills in Europe. Estonia Country Report*.
Environment Food and Rural Affairs Committee (2019) *Plastic food and drink packaging*.
Environmental Audit Committee (2019) *Invasive Species Inquiry launched*. Available at: https://www.parliament.uk/business/committees/committees-a-z/commons-select/environmental-audit-committee/news-parliament-2017/invasive-species-inquiry-launch-17-19/.
Epstein, D., Newhart, M. and Vernon, R. (2014) 'Not by technology alone: The "analog" aspects of online public engagement in policymaking', *Government Information Quarterly,* 31(2), pp. 337-344.
European Parliament (2018) *How many languages are used in Parliament?* Available at: http://www.europarl.europa.eu/news/en/press-room/faq/9/how-many-languages-are-used-in-parliament (Accessed: 10 January 2018.
European Parliament (2020) *Euroscola*. Available at: https://www.europarl.europa.eu/euroscola/en/at_a_glance.html (Accessed: 25 August 2020.
Faralli, S., Stilo, G. and Velardi, P. (2017) 'Automatic acquisition of a taxonomy of microblogs users' interests', *Web Semantics: Science, Services and Agents on the World Wide Web,* 45, pp. 23-40.
Feinerer, I., Hornik, K. and Meyer, D. (2008) 'Text Mining Infrastructure in R', *2008,* 25(5), pp. 54.
Flekova, L., Carpenter, J., Giorgi, S., Ungar, L. and Preoţiuc-Pietro, D. 'Analyzing biases in human perception of user age and gender from text'. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 843-854.
Flekova, L., Preoţiuc-Pietro, D. and Ungar, L. 'Exploring stylistic variation with age and income on twitter'. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 313-319.
Flesch, R. (1948) 'A new readability yardstick', *Journal of applied psychology,* 32(3), pp. 221.
Flinders, M. (2015) 'The General Rejection? Political Disengagement, Disaffected Democrats and 'Doing Politics' Differently', *Parliamentary Affairs,* 68(suppl_1), pp. 241-254.
Forman, G. (2007) 'Feature selection for text classification', *Computational methods of feature selection*, pp. 257-275.

Fredriksson, M. (2015) 'The Pirate Party and the politics of communication', *International Journal of Communication,* 9, pp. 909-924.
Fuka, K. and Hanka, R. 'Feature set reduction for document classification problems'. *IJCAI-01 Workshop: Text Learning: Beyond Supervision*.
Goldfinch, S., Gauld, R. and Herbison, P. (2009) 'The participation divide? political participation, trust in government, and e-government in Australia and New Zealand', *Australian Journal of Public Administration,* 68(3), pp. 333-350.
González-Bailón, S., Borge-Holthoefer, J. and Moreno, Y. (2013) 'Broadcasters and Hidden Influentials in Online Protest Diffusion', *American Behavioral Scientist,* 57(7), pp. 943-965.
Good Things Foundation (2017) *Digital Inclusion*. Available at: https://www.goodthingsfoundation.org/areas-of-work/digital-inclusion (Accessed: 28 November 2017.
Griffiths, T. L. and Steyvers, M. (2004) 'Finding scientific topics', *Proceedings of the National Academy of Sciences,* 101(suppl 1), pp. 5228.
Gurajala, S., White, J. S., Hudson, B., Voter, B. R. and Matthews, J. N. (2016) 'Profile characteristics of fake Twitter accounts', *Big Data & Society,* 3(2), pp. 2053951716674236.
Haegeman, L. (2009) *Thinking syntactically: a guide to argumentation and analysis.* John Wiley & Sons.
Hale, S. A., John, P., Margetts, H. and Yasseri, T. (2018) 'How digital design shapes political participation: A natural experiment with social information', *PloS one,* 13(4), pp. e0196068.
Halfon, R. (2018) 'Repairing the accountability chain between people and Parliament.', *Hansard Society*. Available at: https://www.hansardsociety.org.uk/blog/repairing-the-accountability-chain-between-people-and-parliament (Accessed 11 May 2018.
Hansard Society (2011) *Parliaments and Public Engagement.* London: Hansard Society.
Hansard Society (2017) *Audit of Political Engagement 14*, London. Available at: https://www.hansardsociety.org.uk/research/audit-of-political-engagement.
Hansard Society (2018) *Audit of Political Engagement 15*, London. Available at: https://www.hansardsociety.org.uk/projects/research/audit-of-political-engagement.
Hansard Society 2019. Audit of Political Engagement 16.
Haro-de-Rosario, A., Sáez-Martín, A. and del Carmen Caba-Pérez, M. (2018) 'Using social media to enhance citizen engagement with local government: Twitter or Facebook?', *New Media & Society,* 20(1), pp. 29-49.
Hay, C. (2007) *Why We Hate Politics.* Wiley.
Heutlin, J. (2016) 'The Rise and Fall of the Pirate Party', *New Republic*.
Hilary Spencer 17 May 2018 2018. *RE: Employer Dress Codes.* Type to Helen Jones MP.
Hong, L. and Davison, B. D. (2010) 'Empirical study of topic modelling in twitter', *Proceedings of 1st Workshop on Social Media Analytics*, pp. 1-9.
Hornik, K. and Grün, B. (2011) 'topicmodels: An R package for fitting topic models', *Journal of statistical software,* 40(13), pp. 1-30.
House of Commons 2019a. House of Commons Data Protection information.
House of Commons (2019b) *Mumsnet users in the House of Commons, 2019*. Available at: https://houseofcommons.shorthandstories.com/mumsnet/.
House of Commons Information Office, Office, H.o.C.I. (2010) *The Official Report*.
House of Commons Library, Library, H.o.C. (2018) *E-Petition 200000 "Make British Sign Language part of the National Curriculum"*.
House of Commons Library 2020. General Election 2019: Turnout.

IDEA (2018) *UK Voter Turnout*. Available at: https://www.idea.int/data-tools/question-countries-view/521/137/ctr (Accessed: 13 March 2018.

Inter-Parliamentary Union (2016) *World e-Parliament Report 2016*.

Inter-Parliamentary Union (2018) *World e-Parliament Report 2018*.

Jensen, J. L. (2003) 'Virtual democratic dialogue? Bringing together citizens and politicians', *Information Polity,* 8(1, 2), pp. 29-47.

Jivani, A. G. (2011) 'A comparative study of stemming algorithms', *Int. J. Comp. Tech. Appl,* 2(6), pp. 1930-1938.

Kalampokis, E., Tambouris, E. and Tarabanis, K. 'A Domain Model for eParticipation'. *2008 Third International Conference on Internet and Web Applications and Services*, 8-13 June 2008, 25-30.

Kaufman, L. and Rousseeuw, P. J. (2009) *Finding groups in data: an introduction to cluster analysis.* John Wiley & Sons.

Kelso, A. (2007) 'Parliament and political disengagement: neither waving nor drowning', *The Political Quarterly,* 78(3), pp. 364-373.

Kido, G. S., Igawa, R. A. and Barbon Jr, S. 'Topic modeling based on louvain method in online social networks'. *XII Brazilian Symposium on Information Systems-Information Systems in the Cloud Computing Era*, 353-360.

Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L. and Chissom, B. S. (1975) 'Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel'.

Knispelis, A. 2016. LDA Topic Models. YouTube.

Kolaczyk, E. D. (2009) *Statistical analysis of network data: methods and models.* Springer Science & Business Media.

Krueger, R. A. and Casey, M. A. (2014) *Focus groups: A practical guide for applied research.* Sage publications.

Lee, L. and Young, P. (2013) 'A disengaged Britain? Political interest and participation over 30 years', *Park., A., Bryson, C., Clery, E., Curtice, J. and Phillips, M.(eds.), British Social Attitudes: the 30th Report, London: NatCen.*

Lenihan, D. (2008) *It's More Than Talk: Listen, Learn and Act-a New Model for Public Engagement: the Final Report of the Public Engagement Initiative.* Province of New Brunswick.

Leston-Bandeira, C. (2007) 'The Impact of the Internet on Parliaments: a Legislative Studies Framework', *Parliamentary Affairs,* 60(4), pp. 655-674.

Leston-Bandeira, C. (2012) 'The pursuit of legitimacy as a key driver for public engagement: the European Parliament case', *Parliamentary Affairs,* 67(2), pp. 415-436.

Leston-Bandeira, C. (2016) 'Why symbolic representation frames parliamentary public engagement', *The British Journal of Politics and International Relations,* 18(2), pp. 498-516.

Leston-Bandeira, C. and Thompson, L. (2017) 'Integrating the view of the public into the formal legislative process: public reading stage in the UK House of Commons', *The Journal of Legislative Studies,* 23(4), pp. 508-528.

Leston-Bandeira, C. and Walker, A. (2018) 'Public Engagement', *Exploring Parliament*: Oxford University Press, pp. 304-317.

Liaison Committee (2015) *Building public engagement: Options for developing select committee outreach*, London: House of Commons. Available at: https://publications.parliament.uk/pa/cm201516/cmselect/cmliaisn/470/47002.htm.

Liaison Committee (2019) *The effectiveness and influence of the select committee system*.

Litvinenko, A. 'Social media and perspectives of liquid democracy on the example of political communication of Pirate Party in Germany'. *The Proceedings of the 12th European Conference on e-Government in Barcelona*, 403-408.

Liu, B. (2018) 'Opinion Lexicon'. Available at: https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html#lexicon (Accessed.

Macintosh, A. and Whyte, A. (2008) 'Towards an evaluation framework for eParticipation', *Transforming government: People, process and policy,* 2(1), pp. 16-30.

Maechler, M. (2019) 'Finding Groups in Data'': Cluster Analysis Extended Rousseeuw et'.

Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., Pfetsch, B., Heyer, G., Reber, U., Häussler, T., Schmid-Petri, H. and Adam, S. (2018) 'Applying LDA Topic Modeling in Communication Research: Toward a Valid and Reliable Methodology', *Communication Methods and Measures,* 12(2-3), pp. 93-118.

Maynard, D., Roberts, I., Greenwood, M. A., Rout, D. and Bontcheva, K. (2017) 'A framework for real-time semantic social media analysis', *Web Semantics: Science, Services and Agents on the World Wide Web,* 44, pp. 75-88.

Maynard, D. G. and Greenwood, M. A. 'Who cares about sarcastic tweets? investigating the impact of sarcasm on sentiment analysis'. *LREC 2014 Proceedings*: ELRA.

McCallum, A., Wang, X. and Corrada-Emmanuel, A. (2007) 'Topic and role discovery in social networks with experiments on enron and academic email', *Journal of Artificial Intelligence Research,* 30, pp. 249-272.

Meyer, D. (2012) 'How the German Pirate Party's "Liquid Democracy" Works', *Tech President*. Available at: http://techpresident.com/news/wegov/22154/how-german-pirate-partys-liquid-democracy-works.

Modern Studies Association (2019) *Annual Conference*. Available at: http://www.msa-scotland.net/?page_id=125.

Modernisation Committee (2004) 'Connecting Parliament with the Public (First Report of Session 2003-04 No. HC 368)', *London: The Stationery Office Limited*.

Mohammad, S. M. and Turney, P. D. 'Emotions Evoked by Common Words and Phrases:

Using Mechanical Turk to Create an Emotion Lexicon', *Proceedings of the NAACL-HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Tex*, Los Angeles, California.

Mohammad, S. M. and Turney, P. D. (2013) 'Crowdsourcing a Word-Emotion Association Lexicon', *Computational Intelligence,* 29.

Mosca, L. (2014) 'The Five Star Movement: Exception or Vanguard in Europe?', *The International Spectator,* 49(1), pp. 36-52.

Mutz, D. and Wojcieszak, M. (2009) 'Online Groups and Political Discourse: Do Online Discussion Spaces Facilitate Exposure to Political Disagreement?', *Journal of Communication,* 59(1), pp. 40-56.

Mutz, D. C. (2006) *Hearing the other side: Deliberative versus participatory democracy.* Cambridge University Press.

Natale, S. and Ballatore, A. (2014) 'The web will kill them all: new media, digital utopia, and political struggle in the Italian 5-Star Movement', *Media, Culture & Society,* 36(1), pp. 105-121.

Nesta (2019) *Smarter select committees. Using digital tools to grow digital engagement in Parliament*. Available at: https://media.nesta.org.uk/documents/2019_6_20_nesta_TFoPI_180x260_online_final.pdf.

New York Times (2018) *Italy's Populist Parties Win Approval to Form Government*. nytimes.com. Available at: https://www.nytimes.com/2018/05/31/world/europe/italy-government-populists.html (Accessed: 18 August 2020.

Newman, M. E. and Girvan, M. (2004) 'Finding and evaluating community structure in networks', *Physical review E,* 69(2), pp. 026113.

Nguyen, D., Smith, N. A. and Rosé, C. P. 'Author age prediction from text using linear regression'. *Proceedings of the 5th ACL-HLT workshop on language technology for cultural heritage, social sciences, and humanities*: Association for Computational Linguistics, 115-123.

Nielsen, F. Å. (2011a) 'AFINN'. Available at: http://www2.imm.dtu.dk/pubdb/p.php?6010 (Accessed.

Nielsen, F. Å. (2011b) 'A new ANEW: Evaluation of a word list for sentiment analysis in microblogs', *arXiv preprint arXiv:1103.2903*.

Nikita, M. (2016) 'Select number of topics for LDA model', *Retrieved from*.

Norris, P. (2001) *Digital divide: Civic engagement, information poverty, and the Internet worldwide.* Cambridge University Press.

Norris, P. (2011) *Democratic deficit: Critical citizens revisited.* Cambridge University Press.

Northern Ireland Affairs Committee (2018) *Brezit and Northern Ireland: fisheries*. Available at: https://publications.parliament.uk/pa/cm201719/cmselect/cmniaf/878/878.pdf.

Norton, P. (2005) *Parliament in British Politics. Contemporary Political Studies*: Palgrave Macmillan.

OECD (2009) *Focus on Citizens: Public Engagement for Better Policy and Services.* Paris: OECD Publishing.

Ofcom (2019) *Children and parents: media use and attitudes report 2018*. Available at: https://www.ofcom.org.uk/__data/assets/pdf_file/0024/134907/Children-and-Parents-Media-Use-and-Attitudes-2018.pdf.

Office for National Statistics 2019. United Kingdom population mid-year estimate.

Office for National Statistics (2020) *Internet access - households and individuals, Great Britain: 2020*.

Osman, I. H., Anouze, A. L., Irani, Z., Al-Ayoubi, B., Lee, H., Balcı, A., Medeni, T. D. and Weerakkody, V. (2014) 'COBRA framework to evaluate e-government services: A citizen-centric perspective', *Government information quarterly,* 31(2), pp. 243-256.

OurSpace (2017) *OurSpace: Home*. Available at: http://www.joinourspace.eu/Home/tabid/41/language/en-GB/Default.aspx (Accessed: 13 December 2017.

Oz, M., Zheng, P. and Chen, G. M. (2018) 'Twitter versus Facebook: Comparing incivility, impoliteness, and deliberative attributes', *New Media & Society,* 20(9), pp. 3400-3419.

Papacharissi, Z. (2002) 'The virtual sphere: The internet as a public sphere', *New media & society,* 4(1), pp. 9-27.

Papadopoulos, S., Kompatsiaris, Y., Vakali, A. and Spyridonos, P. (2012) 'Community detection in Social Media', *Data Mining and Knowledge Discovery,* 24(3), pp. 515-554.

Parliament Education Service (2017a) *MP For A Week game*. Available at: http://www.parliament.uk/education/teaching-resources-lesson-plans/mp-for-a-week-game/?referer=parliamentgames (Accessed: 09 November 2017.

Parliament Education Service (2017b) *Outreach*. Available at: http://www.parliament.uk/education/outreach-in-your-school/ (Accessed: 18 December 2017.

Parliament Education Service (2017c) *Teacher resources and lesson plans*. Available at: http://www.parliament.uk/education/teaching-resources-lesson-plans/ (Accessed: 18 December 2017.

Parliament.UK (2017) *Digital Debates*. Available at: https://www.parliament.uk/get-involved/have-your-say-on-laws/digital-debates/?etype=Past (Accessed: 17 November 2017.
Parliament.Uk (2020) *Guidance on giving evidence to a Select Committee of the House of Commons*. Available at: https://www.parliament.uk/get-involved/committees/how-do-i-submit-evidence/commons-witness-guide/.
Parliamentary Digital Service (2017) 'Introducing the PDS content team'. Available at: https://pds.blog.parliament.uk/2017/01/11/introducing-the-pds-content-team%EF%BF%BC/.
Parliamentlive.tv 2018. That this House has considered e-petition 200000 relating to British Sign Language being part of the National Curriculum - During this debate there will be simultaneous British Sign Language interpretation.
Participatedb.com (2019) 'Digital Engagement Catalogue'. Available at: http://www.participatedb.com/.
Parycek, P., Sachs, M., Sedy, F. and Schossböck, J. (2014) *Evaluation of an E-participation Project: Lessons Learned and Success Factors from a Cross-Cultural Perspective.*
Pateman, C. (2012) 'Participatory Democracy Revisited', *Perspectives on Politics,* 10(1), pp. 7-19.
Pennacchiotti, M. and Popescu, A.-M. (2011) 'A Machine Learning Approach to Twitter User Classification', *Icwsm,* 11(1), pp. 281-288.
Persily, N. (2017) 'The 2016 US Election: Can democracy survive the internet?', *Journal of democracy,* 28(2), pp. 63-76.
Petitions Committee (2015) *Role of the Committee*. Parliament.UK. Available at: https://www.parliament.uk/business/committees/committees-a-z/commons-select/petitions-committee/role/.
Petitions Committee, Committee, P. (2016a) *Your Petitions: A year of action*.
Petitions Committee (2016b) *Your petitions: Malcolm Clark's story*. Parliament.uk. Available at: http://www.parliament.uk/business/committees/committees-a-z/commons-select/petitions-committee/news-parliament-2015/your-petitions-malcolm-clarks-story/ (Accessed: 20 November 2017.
Petitions Committee (2017a) *Good afternoon, the response to the petition was from @DExEUgov -the relevant Government department. We are independent from Government* [Twitter]. Available at: https://twitter.com/HoCpetitions/status/915166145296117760 (Accessed: 17 November 2017.
Petitions Committee (2017b) *Make it illegal for a company to require women to wear high heels at work*. petition.parliament.uk: Petitions Committee. Available at: https://petition.parliament.uk/archived/petitions/129823 (Accessed: 29 May 2018.
Petitions Committee (2018a) *Change the laws governing the use of fireworks to include a ban on public use*. Available at: https://petition.parliament.uk/archived/petitions/201947.
Petitions Committee, Committee, P. (2018b) *Debate: family visitor visas*.
Petitions Committee, Committee, P. (2018c) *Online abuse and the experience of disabled people: draft recommendations for consultation*.
Petitions Committee (2020) *How petitions work*. Available at: https://petition.parliament.uk/help (Accessed: 18 August 2020.
Pew Research Centre (2019) *Sizing up twitter users*. Available at: https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users/.
Provost, F. and Fawcett, T. (2013) *Data Science for Business: What you need to know about data mining and data-analytic thinking.* " O'Reilly Media, Inc.".

Pu, P. 'Emotion Analysis in Natural Language'. University of Leeds, 08 December.
Rasmussen, A. (2015) 'Participation in written government consultations in Denmark and the UK: System and actor-level effects', *Government and Opposition,* 50(2), pp. 271-299.
Reform Committee (2009) *Rebuilding the House*. Available at: https://publications.parliament.uk/pa/cm200809/cmselect/cmrefhoc/1117/1117.pdf.
Rieder, B. 'Studying Facebook via data extraction: the Netvizz application'. *Proceedings of the 5th annual ACM web science conference*, 346-355.
Röder, M., Both, A. and Hinneburg, A. 'Exploring the space of topic coherence measures'. 2015, 399-408.
Rose, S., Engel, D., Cramer, N. and Cowley, W. (2010) 'Automatic Keyword Extraction from Individual Documents', pp. 1-20.
Rowe, G. and Frewer, L. J. (2005) 'A Typology of Public Engagement Mechanisms', *Science, Technology, & Human Values,* 30(2), pp. 251-290.
RStudio Inc. (2013) *Easy web applications in R*. Available at: http://www.rstudio.com/shiny/.
Russell, M. A. (2011) *Mining the social web: Analyzing data from Facebook, Twitter, LinkedIn, and other social media sites.* " O'Reilly Media, Inc.".
Sah, P., Leu, S., Cross, P., Hudson, P. and Bansal, S. (2017) 'Unraveling the disease consequences and mechanisms of modular structure in animal social networks', *Proceedings of the National Academy of Sciences of the United States of America,* 114.
Savina, A. (2016) *How Estonia became the most modern digital state in the world* [Medium]. Available at: https://medium.com/@annasavina/how-estonia-became-the-most-modern-digital-state-in-the-world-f777d853aaa6 (Accessed: 20 November 2017.
Sen, A., Rudra, K. and Ghosh, S. 'Extracting situational awareness from microblogs during disaster events'. *Communication Systems and Networks (COMSNETS), 2015 7th International Conference on*: IEEE, 1-6.
Shinyapps.io (2020) *Shinyapps.io by Rstudio*. Available at: https://www.shinyapps.io/.
ShouldertoSoldier (2020) *Shoulder to Soldier*. Available at: https://www.shouldertosoldier.org.uk/.
Sievert, C. and Shirley, K. 'LDAvis: A method for visualizing and interpreting topics'. *Proceedings of the workshop on interactive language learning, visualization, and interfaces*, 63-70.
Silge, J. and Robinson, D. (2016) 'tidytext: Text mining and analysis using tidy data principles in R', *Journal of Open Source Software,* 1(3), pp. 37.
Silge, J. and Robinson, D. (2017) *Text mining with R: A tidy approach.* " O'Reilly Media, Inc.".
Sloan, L., Morgan, J., Burnap, P. and Williams, M. (2015) 'Who tweets? Deriving the demographic characteristics of age, occupation and social class from Twitter user meta-data', *PloS one,* 10(3), pp. e0115545.
Smith, G. (2009) *Democratic innovations: designing institutions for citizen participation.* Cambridge University Press.
Smith, M. A., Rainie, L., Shneiderman, B. and Himelboim, I. (2014) 'Mapping Twitter topic networks: From polarized crowds to community clusters', *Pew Research Center,* 20.
Snowball (2020) *Snowball*. Available at: https://snowballstem.org/.
Social media research group (2016) *Using social media for social research: An introduction*. 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.
Socialfy.pw (2020) *Social Media Marketing Platform*. Available at: https://socialfy.pw/ (Accessed: 24 March 2020.

Spradley, J. P. (2016) *Participant observation.* Waveland Press.

Statista 2019a. Facebook users in the United Kingdom.

Statista (2019b) *Leading countries based on number of Twitter users as of January 2019 (in millions)* Available at: https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/ (Accessed: 11 March 2019.

Statista (2020) *Global digital population as of October 2020*.

Stewart, D. W. and Shamdasani, P. N. (2014) *Focus groups: Theory and practice* Sage publications.

Sunstein, C. R. (2002) 'On a danger of deliberative democracy', *Daedalus,* 131(4), pp. 120-124.

Author (2020) *Tableau Desktop*. Available at: https://www.tableau.com/en-gb/products/desktop.

Taddy, M. 'On estimation and selection for topic models'. *Artificial Intelligence and Statistics*, 1184-1193.

Tamkivi, S. (2014) *Lessons from the World's Most Tech-Savvy Government* [The Atlantic]. Available at: https://www.theatlantic.com/international/archive/2014/01/lessons-from-the-worlds-most-tech-savvy-government/283341/ (Accessed: 20 November 2017.

The Student Room 2020. Official House of Commons,. The Student Room.

Tian, Y., Galery, T., Dulcinati, G., Molimpakis, E. and Sun, C. 'Facebook sentiment: Reactions and emojis'. *Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media*, 11-16.

Tibshirani, R., Walther, G. and Hastie, T. (2001) 'Estimating the number of clusters in a data set via the gap statistic', *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* 63(2), pp. 411-423.

Transport Committee (2019) *Pavement Parking*.

Author (2020) *Tweepy. An easy-to-use Python library for accessing the Twitter API*.

Twitter (2019) *Standard search API*. Available at: https://developer.twitter.com/en/docs/tweets/search/api-reference/get-search-tweets.

Unboxed (2019) *Petition Map*.

Vakulenko, S., Nixon, L. and Lupu, M. (2017) 'Character-based neural embeddings for tweet clustering', *arXiv preprint arXiv:1703.05123*.

van Dijk, J. A. G. M. (2006) 'Digital divide research, achievements and shortcomings', *Poetics,* 34(4), pp. 221-235.

Verma, S., Vieweg, S., Corvey, W. J., Palen, L., Martin, J. H., Palmer, M., Schram, A. and Anderson, K. M. 'Natural Language Processing to the Rescue? Extracting" Situational Awareness" Tweets During Mass Emergency'. *ICWSM*.

Vidgen, B. and Yasseri, T. (2020) 'What, when and where of petitions submitted to the UK government during a time of chaos', *Policy Sciences,* 53(3), pp. 535-557.

Wachs, J., Yasseri, T., Lengyel, B. and Kertész, J. (2019) 'Social capital predicts corruption risk in towns', *Royal Society open science,* 6(4), pp. 182103.

Walker, A. (2012) 'A People's Parliament?', *Parliamentary Affairs,* 65(1), pp. 270-280.

Walker, A., Jurczak, N., Bochel, C. and Leston-Bandeira, C. (2019) 'How Public Engagement Became a Core Part of the House of Commons Select Committees', *Parliamentary Affairs,* 72(4), pp. 965-986.

Wang, C. and Blei, D. M. 'Collaborative topic modeling for recommending scientific articles'. *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*: ACM, 448-456.

Warren, A. M., Sulaiman, A. and Jaafar, N. I. (2014) 'Social media effects on fostering online civic engagement and building citizen trust and trust in institutions', *Government Information Quarterly,* 31(2), pp. 291-301.

Webb, P. (2013) 'Who is willing to participate? Dissatisfied democrats, stealth democrats and populists in the United Kingdom', *European Journal of Political Research,* 52(6), pp. 747-772.

Weber, C. (2013) 'Emotions, Campaigns, and Political Participation', *Political Research Quarterly,* 66(2), pp. 414-428.

Welsh Government, Forestry (2018) *Grey squirrel management action plan for Wales*

Wickham, H., Chang, W. and Wickham, M. H. (2016) 'Package ‘ggplot2’', *Create Elegant Data Visualisations Using the Grammar of Graphics. Version,* 2(1), pp. 1-189.

Wiegand, M., Balahur, A., Roth, B., Klakow, D. and Montoyo, A. 'A survey on the role of negation in sentiment analysis'. *Proceedings of the workshop on negation and speculation in natural language processing*, 60-68.

Author (2019) *UDPipe*.

Wildlifetrusts.org (2019) *Where to see red squirrels*. Available at: https://www.wildlifetrusts.org/where_to_see_red_squirrels (Accessed: 30 September 2019.

Wollebæk, D., Karlsen, R., Steen-Johnsen, K. and Enjolras, B. (2019) 'Anger, Fear, and Echo Chambers: The Emotional Basis for Online Behavior', *Social Media + Society,* 5(2), pp. 2056305119829859.

Yan, X., Guo, J., Lan, Y. and Cheng, X. 'A biterm topic model for short texts'. *Proceedings of the 22nd international conference on World Wide Web*: ACM, 1445-1456.

Yee, N., Bailenson, J. N. and Ducheneaut, N. (2009) 'The Proteus effect: Implications of transformed digital self-representation on online and offline behavior.', *Communications Research,* 36(2), pp. 285-312.

Zalmout, N. and Ghanem, M. 'Multidimensional community detection in Twitter'. *8th International Conference for Internet Technology and Secured Transactions (ICITST-2013)*, 9-12 Dec. 2013, 83-88.

Fin.