

Signal and Noise

Can technology provide a window into the new world of digital politics in the UK?

Alex Krasodomski-Jones

DEMOS

Open Access. Some rights reserved.

As the publisher of this work, Demos wants to encourage the circulation of our work as widely as possible while retaining the copyright. We therefore have an open access policy which enables anyone to access our content online without charge. Anyone can download, save, perform or distribute this work in any format, including translation, without written permission. This is subject to the terms of the Demos licence found at the back of this publication. Its main conditions are:

- Demos and the author(s) are credited
- This summary and the address www.demos.co.uk are displayed
- The text is not altered and is used in full
- The work is not resold
- A copy of the work or link to its use online is sent to Demos.

You are welcome to ask for permission to use this work for purposes other than those covered by the licence. Demos gratefully acknowledges the work of Creative Commons in inspiring our approach to copyright. To find out more go to www.creativecommons.org



PARTNERS CREDITS

This paper was written by the Centre for the Analysis of Social Media (CASM) at Demos in partnership with BCS, The Chartered Institute for IT.



BCS, The Chartered Institute for IT, is committed to making IT good for society. We use the power of our network to bring about positive, tangible change. We champion the global IT profession and the interests of individuals, engaged in that profession, for the benefit of all.

<http://www.bcs.org> @BCS

Published by Demos May 2017
© Demos. Some rights reserved.

Unit 1, 2-3 Mill Street
London SE1 2BD

T: 020 7367 4200

hello@demos.co.uk
www.demos.co.uk

CONTENTS

Foreword	4
Executive Summary	6
The State of Digital Politics	8
Research Questions	15
Recommendations	34
Annex 1: Methodology	38
Annex 2: The Dashboards	44
Annex 3: Technology overview	52

FOREWORD

The way that many of us live our lives online nowadays is naturally spilling over into the way people engage with politics and with politicians.

Accompanying the rise of online campaigns, e-petitions and political memes, the internet is shifting the ways in which citizens engage with their elected representatives. This shift is as fundamental as it was with the advent of radio or television.

In their attempts to cope, some MPs try to avoid digital communications altogether. Others struggle to manage the immense volume of direct public engagement made possible by social media channels.

Both in the UK and abroad, we are increasingly seeing citizens being targeted by social media messaging aiming to influence their voting behaviour, something that presents potentially lasting impacts on our democratic state, society and social cohesion.

Against that backdrop, and with a General Election imminent, this report is timely. MPs will be dutifully engaging as many of their constituents as they can over the coming weeks, both face to face and online. However, in doing so they will be laying themselves open to both praise and the unkind attacks on their public service, lifestyles and personalities.

What is clear though, is that the anonymous and 'safe distance' nature of social media platforms allows such abuse to be handed out far less respectfully than it would usually be if delivered face-to-face.

Politicians of all parties have recently been highlighting this issue and its negative impact on the political process. Campaigns like Reclaim the Internet, internet entrepreneurs like Baroness Martha Lane-Fox and World Wide Web creator Sir Tim Berners-Lee have individually been asking how social media platforms could rethink their approach to encourage the sharing of more constructive thoughts.

Collectively, we need to ask ourselves, if we value a free and fully-functioning digital function, 'Who guards our civil discourse?'

Social media companies are not the enemy here; their platforms are powerful communication tools being used by citizens and politicians to perform a specific societal function for which they were not designed.

But this issue is not going to either diminish or go away; the number of people communicating with their MPs online is not going to reduce. This report suggests options for a different perspective and approach to the current situation.

The Palace of Westminster, where our elected representatives meet and interact with each other, is in a state of disrepair bad enough that the functioning of day-to-day politics is increasingly difficult. As a result, government is being urged to spend billions of pounds to fund its refurbishment.

Yet, as this report identifies, political engagement online, where citizens are increasingly interacting with their elected representatives, is not functioning in a manageable or societally beneficial way.

That is why, working together and for a tiny fraction of the cost of refurbishing Parliament, we could and should be encouraging better political discourse online.

To address this, we at BCS are calling for a cross-party allegiance to work with us and existing social media platforms to improve their offerings, and establish a purpose-built platform to facilitate meaningful and effective political engagement online.

Online political engagement is here to stay, and issues around how well it is serving our political process will only increase with time. The time to give proper consideration to how the situation can be improved, making IT better for society, is now!

David Evans

Director of Policy & Community

BCS, The Chartered Institute for IT

EXECUTIVE SUMMARY

Society, and by default, our elected representatives are increasingly influenced by social media. Whilst this potentially presents vast new opportunities for MPs, it also presents fundamental challenges.

Social media allows MPs to engage, or to be engaged with the electorate in that it allows them access to new ideas, new people, to listen to constituents, to gauge the public mood. But it can also leave them feeling – like many citizens - overwhelmed by digital information. Email campaigns, tweets, Facebook posts: they are unable to make sense of the digital 'noise', and, as a result, often feel unable to respond. This lack of response risks undermining confidence in MPs, who, in the eyes of their electorate, may as a result appear out of touch or unresponsive.

Social media also creates measures of opinion organised by themes that can be difficult for MPs to understand or value, relative to more traditional measures of influence or representation. For example: whether a large Facebook group with lots of members genuinely reflects popular opinion.

Perhaps more fundamentally, although less obviously, online behaviour also creates new expectations: about what people expect of their elected representatives and what MPs might expect in return.

This short paper examines one aspect of this very broad challenge: whether technology can (and should) help MPs make better use of social media by collecting Twitter data and subjecting it to a series of analyses.

We conclude from our analysis that:

- A large amount of politics and political conversation is now taking place online, creating a new 'digital commons.'
- The data created by this digital commons is too large for an MP to manage, and technological solutions are necessary
- Technology is effective in handling the data that makes up this new digital commons

There is an urgent need for technologists to work with political institutions in the UK to make this technology available and usable, and for political institutions and politicians to work alongside, support and encourage its development.

Accordingly, we make three urgent requests:

1. For MPs (and activists) to work collectively to create a new culture of politics online.
2. For existing social media platforms to adapt and become more suited to the task of conducting our politics through them.
3. For technologists and policy makers to work together to assist MPs in developing their own set of bespoke tools and techniques that can aid their functioning as elected representatives, and set out what components such a piece of technology should include.

It is widely acknowledged that online discourses take place under a different set of norms to those online. This was emphasised time and again in discussions with MPs and other members of political parties, who repeated the line that much of what they found unpleasant or counterproductive in online politics simply didn't happen offline. This is part of a wider story: long-established cultures, expectations and pressures around the way we communicate have not taken root online.

For digital politics to be a success, effort will have to go towards creating a digital culture in which it can take place.

As the report shows, a huge amount of political activity takes place online. UK politics has had to adapt to digital tools that are by no means tailored to the structures, expectations and traditions of our democracy. We must carefully identify the positive examples of digital politics in the UK and celebrate them, and identify the negative and call them out. Where possible, politicians should work closely with the major technology companies to identify improvements in existing platforms would improve the democratic process in the UK.

For many reasons, some of which are outlined below, existing platforms are not tailored to politics. As we move towards an increasingly digital society, and expectations of our politics and democracy change to match this, we must ensure the technology we conduct our politics through is fit for the task. This is likely to require a dedicated platform for the digital democracy in the UK.

OUTLINE

The paper begins with an overview of the state of digital politics in the United Kingdom. The second section looks at what British politics might gain from new or improved digital channels, the risks that the change might pose, and the role technology might play in that change. The third section sets out the three research questions. The final section sets out the recommendations. There is a detailed methodology, and a series of potential 'data dashboards' in the Annexes.

This report has been written in collaboration with BCS, The Chartered Institute for IT. Both our organisations are keenly interested in the increasing influence of technology and the digital world on British politics and society, and in ensuring that we are making the most of the opportunities it might provide.

THE STATE OF DIGITAL POLITICS

Introduction

That the vast majority of us are digital citizens is now beyond question.ⁱ Nearly nine in ten adults in the UK are internet users, increasing to 99 percent for those aged 18-24.ⁱⁱ When we are online, we use social networks: 88 percent of internet users are members of at least one social network, per minute spent the most popular online activity. Facebook counts 31 million people in Britain among its users, while Instagram has 14 million and Twitter 13 million.ⁱⁱⁱ

Increasing use of the internet has brought with it a growing expectation that politics ought to do the same. A 2014 study by Ofcom found that one in five people (19 percent) had contacted local councillors online, over a third (35 percent) had signed a petition and nearly half (44 percent) had used the internet to find information about political campaigns.^{iv}

In response, politicians and political institutions have logged on. 86 percent of British MPs are on Twitter, sending tens of thousands of tweets a year, and many are active across multiple platforms. The institutions of government have also embraced the online world. All major departments have social media accounts, as do all local councils, and digital publication of parliamentary information, media and Hansard records were promised by the 2015 parliamentary digital democracy commission and delivered in 2016.^{v vi}

The foundations of digital politics have been laid, and the last decade has witnessed a building of a political culture online. This was illustrated in the results of a survey commissioned by Demos of Ipsos Mori in 2015.^{vii} Half of users (51 percent) surveyed had turned to social media for politics, with nearly three-quarters of those felt it boosted feelings of political engagement (72 percent), with young people particularly well represented. When survey respondents were asked whether they felt social media improved the democratic process, 42 percent agreed or strongly agreed (20 percent disagree/strongly disagree) and 39 percent said they felt more likely to vote as a result (19 percent disagree/strongly disagree).

Digital politics is not limited to social media messaging, nor is it limited to a politics bound by the principles of representative democracy. Political groups far larger than any of the UK political parties by membership have sprung up: Britain First, a far-right political movement, boasts over a million and a half subscribers.^{viii} From media sites and blogs to crowdfunding platforms and petitioning sites, the ecosystem is a fragmented one. There are five UK specific petitioning sites alone that frequently pull in millions of signatories. Dozens of technological solutions are being trialled: crowdsourcing proposals, distributed decision-making and blockchain applications.^{ix}

Yet as the digital world plays a growing role in the lives of an increasing number of people, the opposite might be said for our political institutions. As interest and involvement in all things digital has increased, the last twenty years have equally been characterised by falling trust, participation and enthusiasm for political institutions. It is on this background that we understand the possibilities and pitfalls presented to politicians by technology.

2016 was a bruising culmination to decades of difficulty for establishment politics. Political participation is falling. Membership of political parties has declined sharply. Electoral turnout is in decline. Since the millennium, turnout in general elections has averaged 62 percent. Those who have taken to the online world are those least likely to participate in politics. In 2015 just 43 percent of 18-24 year olds turned out to vote. Trust is also falling. Politicians are consistently the least trusted profession in the UK. A survey in 2015 found just one in five people trusted politicians.^x

The immediate answer is obvious: to engage with the young and disaffected, politicians need to join them online. But this may be a dangerous oversimplification. The growth of politics online has not been without pain. The economic model of digital media appears to favour

the sensational and the attention-grabbing, the short and the simple. Political supporters are at risk of communicating in 'echo chambers': confirmatory online spaces offering the illusion of plurality of opinion.^{xi} A number of MPs have spoken out against racist, Islamophobic, anti-Semitic and misogynist online abuse they have received.

More worrying still, while some surveys have shown the power of the internet in bringing people closer to politics, data has also shown that old problems persist. Online patterns in political engagement closely match offline patterns: once online, young people are *still* less likely to engage with politics than the more traditionally politically-active generations. This suggests that simply 'getting online' is not the answer: a digital democracy will require fresh methods and new tools if it is going to reassert its relevance in the internet age.

Possibilities

The digital revolution has disrupted politics, but it could enhance democracy. It offers new opportunities for MPs to reconnect with voters, and vice versa. From the perspective of a representative, digital technology opens up many new opportunities to improve the directness, speed and relationship with the people they represent.

New ways to 'listen'

One of the recurring complaints in modern representative democracies is that politicians are 'out of touch' or 'do not listen' to ordinary voters' concerns. Ipsos Mori polling from September 2015 found 64 percent of respondents felt David Cameron was "out of touch with ordinary people", while the leader of the Labour Party, Jeremy Corbyn, was deemed out of touch by 39 percent of respondents (rising to 44 percent in 2016). The EIU's 2016 report "The Revenge of the Deplorables" refers to a "popular revolt in 2016 against political elites who are perceived by many to be out of touch and failing to represent the interests of ordinary people".

Social media offers a new way for politicians to better connect. After all, the UK telecoms regulator Ofcom has found that British adults spend an average of eight hours and 41 minutes a day on media devices, against an average night's sleep of eight hours and 21 minutes.

The result has been a dramatic explosion in public opinion data generated by citizens that can potentially be used by MPs to better understand constituents' (or the general public's) concerns, frustrations or interests more quickly and easily.

Digital technology allows people to find myriad new ways to express their political views publicly, outside of formal political spaces. Every day there are millions of conversations about political issues in new digital spaces: on Twitter timelines, Facebook newsfeeds, comment threads, blogs and videos. This new 'digital commons' reflects the hopes, views and beliefs of citizens. Social media offers a new way for MPs to better understand voters' concerns

New ways to reach out

The way that social media is changing political engagement and activism is a relatively new area of research, but recent studies in the UK indicate that the internet has become a vital new avenue of political activism. One recent, representative poll of 1000 British social media users revealed that over half of them either sent or received political material on social media over the last three months, totalling around 11 million people overall. This is greater than the number of social media users who reported participating in politics or activism offline. What's more, this activity is proactive, not passive. In the three months to May 2015, as many people (40 per cent) contributed political content on social media as received it (38 per cent).

This suggests that online activity is one way for people to get involved in politics. Indeed, the survey above found that the majority of Brits surveyed felt that social media improved the democratic process by encouraging more open discussion and greater access to debate. They better understand the issues and what the parties stood for; they feel more engaged in the political debate; and they are more likely to vote. If extrapolated to all 23 million social media users in the UK, it would mean that over 4 million people felt they understood the general election campaign better, and were more likely to vote as a result of political activity on social media. While this is by no means a new trend – writers such as Manuel Castells and Clay Shirky have been arguing this for some time – it remains underdeveloped within the major political parties. Social media offers a way for MPs to reach out to voters and possible supporters.

New ways to create interest groups

Political activism is increasingly single interest based: growing numbers of people prefer to mobilise around specific events or issues rather than join a party. The last two years has seen major single-interest movements mobilising quickly with the assistance of digital technology (the election of Jeremy Corbyn is one example of this, although it morphed into a much larger movement). For MPs or elected representatives who in addition to their constituency duty, campaign on or are interested in specific thematic areas, social media is an important new channel to these groups. Social media is a new place for political movements, which are often not closely associated with individual parties.

The digital challenge

Each of the opportunities set out above also creates problems. None of the above are easy to realise. While there are exciting new prospects for representative democracies, digital technology is also an extremely disruptive force.

New types of opinion

Traditional methods of understanding public opinion, such as representative polls or direct interaction with constituents, have been supplemented – although certainly not entirely replaced – by new online expressions of public opinion such as online lobbying groups, e-petitions, Facebook pages and Twitter campaigns. But it's not clear how useful these new sources actually are. No social media trend is nakedly a proxy for public opinion. Only certain parts of society use social media platforms, and usually a small number of 'power users' dominate even this conversation. What's more, online trends and virals can be engineered. A new breed of 'guerrilla' marketing agency has sprung up with promised expertise in 'seeding' virals on the internet. It's a difficult art, but a small number of skilful viral marketers can make a carefully engineered campaign look like an organic and spontaneous outburst of public sentiment.

For democracies to function, politicians have to avoid capture by special interests. New technology can make this more difficult when the intensity and scale of online debates make it hard for political actors to distinguish the signal from the noise. Social media data of the kind we present above is particularly high in volume and complexity, making it difficult for political actors to gauge when online data reflects broader public opinion, or even just the views of most online users. A seemingly

large debate with many supporters of a particular position can instead be the result of a campaign by a PR firm, or a small number of angry people who are very active online. These effects of manipulation and amplification have always affected debates, but they are harder to discern in online activity, so political actors need better ways of telling which communications really reflect the views of many citizens. How can MPs understand and correctly value new forms of public opinion expressed online; and how to do so without getting 'captured'?

New expectations

Voters have a growing sense that political parties and law-making are out of touch, but not that politics is irrelevant. This is understandable given the changes in the way we interact with other parts of our lives. Online life is instant, transparent, easy and connected – while politics is often slow, laborious and secretive.

This is especially difficult for MPs, who often receive enormous amounts of digital communication from their constituents, and struggle to respond. The danger is that MPs are incapable of meaningfully responding to even just a fraction of the online correspondence they receive, and the result is that citizens feel they are being ignored, while MPs simultaneously feel overwhelmed. The introduction of exciting new reforms can do more harm than good if they ultimately fail to deliver a better quality of engagement for citizens. How to ensure that people don't feel let down by online interaction?

New technologies

There is a growing industry in various forms of 'data analytics' or 'big data analytics', which allow people to collect and analyse large volumes of data generated online. But this brings with it a whole new suite of skills and techniques that are alien to most MPs, such as machine learning algorithms, application programming interfaces, and data visualisation. These are increasingly the tools required to make sense of the digital world; and yet remain the preserve of a very small number of specialists. What are the basic skills required for MPs to better make sense of the digital world?

New legal and ethical responsibilities

Several political parties have already started to use social media to better understand citizens, and target voters in new sophisticated ways through social media platforms. The potential for misuse of social media

data – especially twinned with analytics software – is considerable, and could even damage confidence in the electoral system. How can MPs best use social media in an ethical and legal way?

Conclusions

The dilemma of social media is that it offers an unprecedented opportunity to help reinvigorate representative – or even forms of direct – democracy, while simultaneously presents problems of capture, misuse, and data-deluge that can make MPs work even more difficult.

This paper examines one aspect of this problem: how far existing technologies could be used to provide useful analysis of social media for MPs and how that might support more effective use of digital channels. In the next chapter, we present the case studies which throw further light on this problem.

RESEARCH QUESTIONS

This paper tests how far existing technologies could be used to provide useful analysis of social media for MPs and how that might support more effective use of digital channels.

To do that, we asked three key research questions designed to determine whether or not social media is potentially a useful and meaningful place for them to improve their responsibilities as elected representatives.

Research question 1: What sort of volume and type of political conversation is taking place online?

This involved several questions:

- What is the volume of tweets being sent to MPs?
- How many users are involved in the conversation?
- How problematic is the volume of conversation for an MP trying to interpret or respond to it?
- What information can be used to quickly prioritise or filter the data?

Research question 2: Can social media data be filtered for 'sentiment' or 'abuse'?

This involved several questions:

- Where can content analysis be used to aid MPs' interpretation of the digital commons?
- Can tweets be classified into 'boos' and 'cheers' for MPs?
- Can abusive material be filtered out?

Research question 3: How well did the technology perform?

While there are lots of technology solutions available on the market for social media analytics, their usefulness and effectiveness can vary. We ran a series of tests to determine how robust social media data analytics are. This involved the following questions:

- Could algorithms identify the categories of meaning demanded by the analyst in the data?
- How effectively were they able to make distinctions between the categories?

- Where were they most effective, and where least?
- What are the long-term risks to using algorithms to analyse language at scale?

A full methodological breakdown is available in the annex. As with much existing literature, Twitter is used here as the source of data as it is the only major social media network which makes a good cross-section of its public data available to researchers, and is home to much political discussion in the United Kingdom.

The dashboards were built using Qlikview and Qlik Sense, visual analytics platforms that allowed the researchers to build bespoke interfaces into the data and share them online.

An opportunity to field-test the dashboard prototypes emerged during the project, and the technology was shared with four MPs and informal feedback was collected through a team of digital support assistants. Further feedback on the findings were gathered during three presentations of the dashboards at the 2016 Labour, Conservative and Scottish National Party conferences.

RESEARCH QUESTION 1: WHAT SORT OF VOLUME AND TYPE OF POLITICAL CONVERSATION IS TAKING PLACE ONLINE?

The first research question used the technology to characterise the scale of political conversations on the platform. In conversations held with MPs and their parliamentary assistants, a common theme was that many were keen on using social media but overwhelmed by the number of messages they might have to respond to. Capturing the scale of the messaging was the first task.

Between 9th May – 18th August 2016, we used Twitter's public API to collect all tweets sent to or from a UK Member of Parliament. The decision was made to focus on tweets mentioning an MP. This excluded political conversation that would not have, in theory, sent an alert to a sitting Member of Parliament.

In total, this dataset contained 11.4 million tweets at an average of 110,000 a day and from 891,000 unique Twitter accounts. At its peak, on the day of the referendum result, over 400,000 tweets were sent to MPs. During our collection period, MPs sent 25,000 tweets, or one for every 15 sent at them.

This average masks a wide variation in each MP's experience of the platform. At one end of the scale, three MPs received, on average, over 10,000 messages a day. 18 received, on average, fewer than five a day.

Variation also occurred within an MP's timeline: it was a common occurrence for an MP to suddenly receive ten times their normal number of mentions. This can be loosely illustrated by identifying the number of tweets an MP received in a single day against their average over the collection period, as in the image below: each MP's row turns orange on a day they received ten times, red when they received 20 times and black when they received 30 times their daily average.

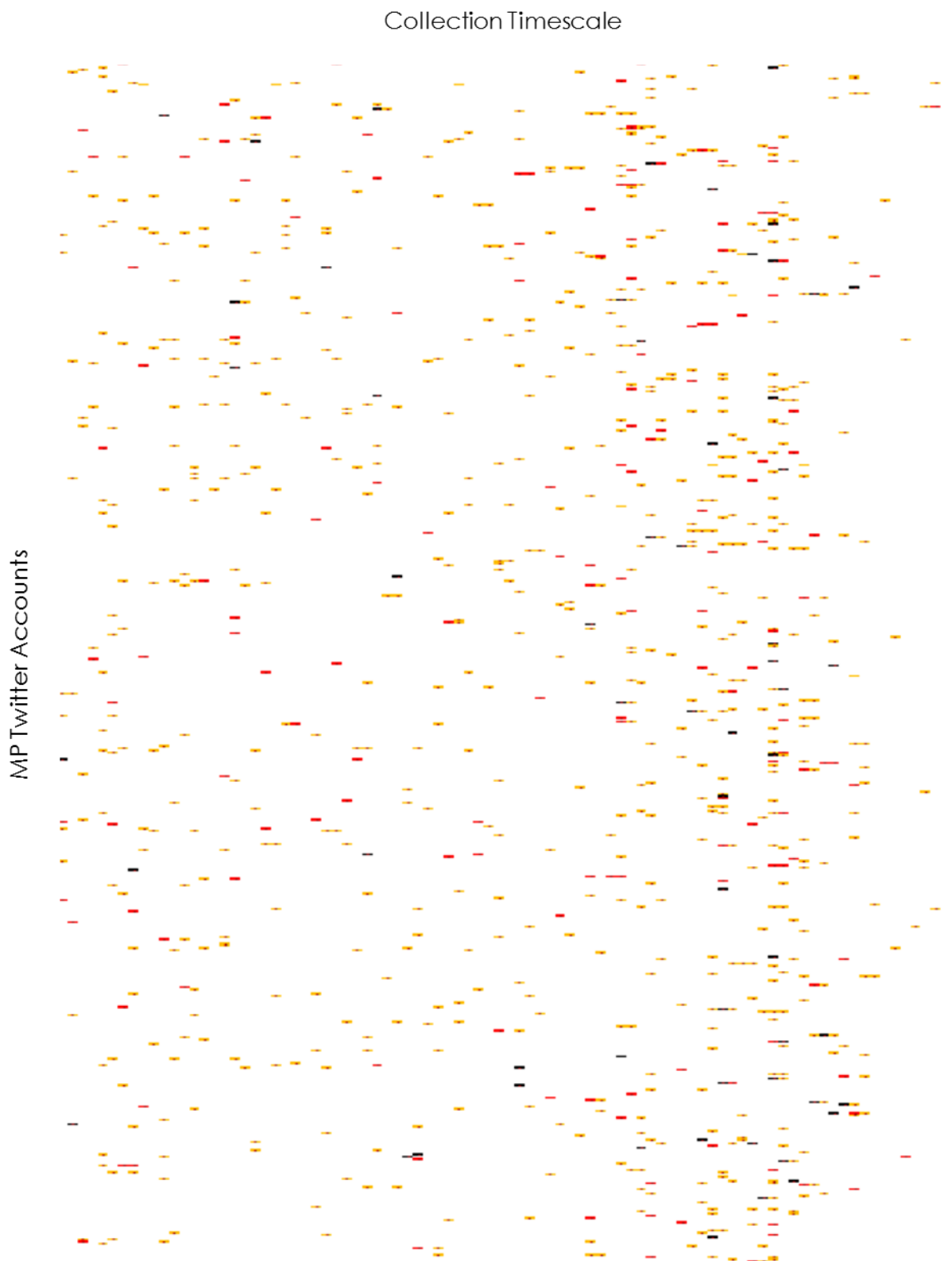


Figure 1: Tweets mentioning MPs sent over the collection period.

The randomness of the plot helps illustrate how from one day to the next, an MP may unexpectedly find themselves under a barrage of Twitter messages before the platform falls silent for a period.

The unfiltered dataset contained messages from nearly a million different Twitter users. Of these, approximately 60 percent were from accounts algorithmically determined to be male, 30 percent female and 10 percent from an organisation or institution.

Automatic annotation of this data provides the first meaningful way to filter it. This could be done in a number of different ways:

1. Language – Automatically annotating the language of a message and filtering it if non-English. This data is contained in the Twitter metadata. When tested this reduced the number of tweets in the dataset by approximately 9 percent.
2. Geography – Automatically locating Twitter users outside of the United Kingdom and excluding them. Alternatively, identifying Twitter users inside the United Kingdom and excluding the remainder. Owing to the accuracy of the classifiers, excluding those definitely outside the UK was felt a better approach than attempting to include only those in the UK. A full methodological write-up of the geographic annotation process is contained in the technical annex.
3. Specific Metadata – It would be possible to, for instance, remove all messages shared containing a link to an external website, or containing a specific hashtag. Analysts were able to identify a number of automated accounts sending hundreds of tweets a day to MPs by their hashtag use, and in this case the tweets were filtered out of the dataset.
4. Duplicate Messages – Analysts investigated the impact of de-duplicating the dataset to prevent a reader encountering multiple messages that were identical. In tests this removed around 55 percent of messages.

Filtering the dataset to English language tweets, not geolocated to outside the UK and not containing links to external sites reduced the dataset to 3.7m tweets, or an average of 36,000 tweets a day. This could be reduced further, to 1.7m (or just over a tenth of the unfiltered data) by removing duplicate messages.

Three things are clearly indicated by the figures reported above. First, that there is an enormous amount of political activity taking place on social media platforms in the UK, and that hundreds of thousands of users are turning to Twitter to contact an MP directly. Second, that the volume of activity is for some MPs, and at some times, too great to be understood by a single user, particularly when an MP's profile or media prominence triggers a 'spike' in activity. Third, that filtering or prioritising Twitter data based on pre-existing or automatically added annotation could improve an MP's experience of the platform.

RESEARCH QUESTION 2: CAN SOCIAL MEDIA DATA BE FILTERED FOR 'SENTIMENT' OR 'ABUSE'?

The second research question investigated the use of content analysis. Where metadata analysis turns on using existing data to characterise a tweet, such as the time it was sent, content analysis adds a layer of algorithmic analysis to label a tweet based on what the user has said.

Two algorithms were tested. First, a 'boos and cheers' algorithm to measure the ways in which Twitter users responded to MPs. To do this, analysts compared positive and negative feedback with other measures to identify whether certain groups of users, of MPs, or of discussion topics and real-world events impacted the type of messages being sent to MPs. Second, an 'abuse' algorithm, aimed at identifying abusive messages, highlighted by MPs and others as one of the biggest challenges posed by the rise of digital platforms as venues for politics.

This analysis used the filtered dataset noted above, and contained 3.7m tweets sent between 9th May – 18th August 2016.

Boos and Cheers

All tweets were categorised as either a 'boo' (a disagreement, an accusation or an insult), a 'cheer' (praise, agreement or support), or neutral (everything else, such as sharing news, quoting the MP or commenting on a wider issue).

In total, 1.8 million messages containing either a boo or a cheer were sent during the period (48 percent of the total dataset). Of these, 730,000 were cheers and one million were boos. The remainder were determined to be neutral. There was a wide variance in the proportion of boos and cheers received by the parties' MPs during the collection period.

Table 1: Cheers and Boos by Party

Party	Total Boos/Cheers	% Cheers	% Boos
Conservative	1023726	32.5%	67.5%
Labour	730864	50.8%	49.2%
SNP	50044	61.6%	38.4%
Lib Dem	15732	62.3%	37.7%
UKIP	12069	24.5%	75.5%
Green	9836	74.0%	26.0%
DUP	3330	49.2%	50.8%
Independent	2629	41.3%	58.7%
Plaid Cymru	1731	69.0%	31.0%
UUP	655	59.4%	40.6%
SDLP	422	85.5%	14.5%
Sinn Féin	182	79.1%	20.9%

UKIP's then single MP, Douglas Carswell, and Conservative MPs proportionally received the highest proportion of booing on Twitter, with smaller parties receiving the highest proportion of cheers. However, this masks great variation across MPs within the larger parties. Within the Conservative party, for instance, seven MPs received critical or unsupportive tweets 90 percent of time, while nine others received less than 10 percent.

Analysts looked for correlation between MPs' use of the platform and the feedback they received. A user's follower count is a useful an indicator for how well-known they are, both on and off Twitter. There was little correlation between this and the number of boos or cheers they received (0.37). Tweets sent from the account were not correlated with boos or cheers received (0.03), nor was the number of replies or retweets that the user sent as a proportion of their tweets (0.11 and -0.95 respectively). Of the top ten most cheered MPs, four were women and six were men.

Discussions of alleged Conservative electoral misconduct were the angriest (92.8 percent boos), but the weekly Prime Minister's Questions also prompted a lot of criticism (81.5 percent boos). The hashtags associated with the opposing Leave and Remain sides debating the EU referendum were strongly contrasting. Four in every five emotive tweets on #Remain were negative, compared with #TakeControl and #VoteLeave with 97.7 percent and 61.3 percent cheers respectively.

The hashtags supporting Andrea Leadsom MP's leadership bid were among the most positive: #LeadsomForLeader, #AndreaLeadsomForLeader and #Freshstart were all almost entirely supportive. A similar pattern can be seen on the hashtags supporting Jeremy Corbyn's leadership campaign. This suggests that in some cases, Twitter is used by UK users as a partisan tool for campaigning through which supporters can team up and cheer their candidate.

The hashtags belonging to television and media are also interesting: the two televised debates (#ITVEURef and #BBCDebate) were watched by Twitter users largely cheering their candidates on. By contrast, the established BBC political television programmes (Question Time and Newsnight) were greeted with boos. When viewed alongside the scores for the EU referendum hashtags, it illustrates the other side to UK political Twitter: contested hashtags where users are cheering and booing the political events they are watching unfold.

To further investigate the reason users in the UK take to Twitter to boo or to cheer, analysts manually coded 300 tweets (150 cheers and 150 boos) to identify the subject of the message and categorise them based on the themes present in the data.

The categories chosen were:

EU Campaigning (Boos and Cheers)

Tweets sent booing or cheering an MP's actions in the EU referendum debate.

had a fabulous time yesterday in lichfield campaigning for #brexit with @XXX @vote_leave

one minute @XXX is a friend of terrorists but now his @XXX best mate.. so still a friend of terrorists then #voteleave

Personal/Personal Behaviour (Boos and Cheers)

Tweets praising or criticising an MP for their personal actions or conduct outside of any specific political issue.

@XXX good to see a sophisticated lady, a powerful voice in the #voteleave campaign. the glamour is a bonus lol!

@XXX because i said she was undemocratic blocking people and hiding their comments. so she blocked me!

Political Behaviour (Boos and Cheers)

Tweets cheering or booing an MP based on political actions they have taken.

@XXX good to hear support for music, drama and art as core components of child's entitlement curriculum - and vital for the economy.

tory £30pw disability benefit cut @XXX on sugar tax: far better if children were simply advised to move about more and eat less

Insult (Boos)

Unqualified insult not clearly related to political or personal actions by an MP.

@XXX and she's got a face like a rivet catchers glove, which goes well with her extra large gub.

@XXX i could probably buy you over with a packet of skittles and pocket change in order for you to disregard your beliefs.

Party (Boos)

Tweets critical of the MP's party as a whole.

@XXX cameron, bliar and brown need to be investigated under electoral fraud and our corrupt judiciary can attempt to clean up

@XXX #whoops #toryelectionfraud

Event (Cheers)

Tweets celebrating, thanking or promoting an MP's attendance at an organised event.

thanks @XXX & @XXX @tentthinktank! great secretariat for new #parliamentary #entrepreneurs group

we would like to thank @XXX for taking time out of her schedule to come see us and talk about our promotion

The results are shown below.

Table 2 & 3: Tweets classified as 'Cheers' and 'Boos' (by Type)

Cheers		
Theme	# Tweets	%
Event	53	35%
Political Behaviour	29	19%
EU Campaign	23	15%
Personal/Personal Behaviour	18	12%
Misclassified/Other	27	18%

Boos		
Theme	# Tweets	%
Personal/Personal Behaviour	37	25%
Insult	31	21%
EU Campaign	20	13%
Political Behaviour	17	11%
Party	17	11%
Misclassified/Other	28	19%

Event attendance by MPs drew the most number of cheers in the sample by percentage (35 percent). This may reflect Twitter's role in galvanizing the vanguard: a tool for encouraging existing supporters. The remaining supportive tweets that could be coded were either political or politics-related in the EU referendum campaign. The personal behaviour of MPs was least likely to draw praise in this sample (12 percent).

This is in stark contrast to the booing tweets, where criticism of the personal behaviour of an MP (25 percent) and general insults (21 percent) make up nearly half the sampled tweets.

The percentage of misclassified/other tweets is lower than expected, suggesting the classifier may be performing better than the F-score noted in the appendix.

Abuse

The analytical pipeline built to identify abusive tweets, described in the methodology section below, was found to be one of the most difficult to build on the dataset. Abuse can be extremely subjective, and there was a lot of disagreement among analysts and observers as to what ought to be classed as abusive. It was decided that forceful disagreement, for instance, should not be classified as abusive, though personal insults should. An illustration of how difficult it is to get agreement on this emerged when the decisions made by the algorithm were presented to audiences. On a sample of just five tweets, some observers felt all should be classified as abusive and others that none should be.

This has important ramifications for technology that monitors the social internet: clear and transparent definitions for the categories of analysis used are important in ensuring users understand why a tweet has been classified one way or another.

188,000 tweets collected over the period were identified as abusive (5 percent of the total dataset). These tweets were sent or retweeted by 130,000 different users. A third of those tweets came from accounts classified as belonging to women, 59 percent from men (a 3 percent increase on the average for all tweets from women and no change from men).

Abuse peaked on two days: the 24th of June, the day of the EU referendum results (7,300 tweets) and on the 30th June, the day Boris Johnson MP pulled out of the Conservative leadership race (5,000 tweets).

On average, one in 20 tweets and retweets mentioning MPs were classified as abusive (5.1 percent). However, as with boos and cheers, MPs in the UK have markedly different experiences. Six MPs received one abusive tweet in every ten, while 34 received none at all. Eight of the MPs receiving the most abuse were male, two were female. There was little correlation between the follower count for an MP and the abuse they received (0.31).

A qualitative analysis of the 25 most frequently abused MPs gives some indication of why they were targeted. The list includes the leaders and deputy-leaders and challenges for the leadership in both the Conservative and Labour parties, as well as prominent campaigners for and against leaving the European Union (21 of 25). Two were high-profile

politicians linked to the NHS, and two were female MPs who spoke out against abuse online.

The role of the EU referendum debate in provoking abuse is supported by an analysis of the hashtags linked to abusive tweets. The ten hashtags most frequently co-occurring with abusive tweets are shown below.

Table 4: Most frequently occurring Hashtags (Abusive Tweets)

Hashtag	Abusive Tweets
#brexit	4870
#borisjohnson	3694
#voteleave	2463
#reclaimtheinternet	1554
#strongerin	950
#leaveeu	922
#euref	803
#pmqs	670
#remain	655
#bbcdebate	584

Eight of the ten are related to the EU referendum. #ReclaimTheInternet was a campaign started during the collection period raising awareness of abuse online. #PMQs is the hashtag used to comment on the weekly Prime Minister's Questions.

The peaks in abuse, the targeted MPs and the hashtags all indicate that during the EU referendum, intensified levels of abuse were received by the MPs involved. It suggests that high-profile events and media events, particularly when divisive, provoke increased abuse for the politicians at their forefront.

A closer look at the data was performed through a second qualitative study of a sample of abusive tweets. 150 abusive tweets were coded by an analyst into four broad categories.

European Referendum

Abusive tweets targeted at an MP specifically regarding their role in the debate surrounding the European Referendum.

@XXX eu bought and paid for, self serving, turncoat parasite.

@XXX i have no faith in you, your agenda, your policies, your party.. you talk utter nonsense, scaremongering won't work #brexit

Political

Abusive tweets in response to a politician's actions or stance on a political issue outside of the European Referendum.

@XXX you two faced fascist. so you're the blairite shafted member. we know of you #corbyn4pm

@XXX "working families" something you know fuck all about you cock wamble

Personal

Non-political abuse aimed at a target's appearance, religion, ethnicity, gender or personality.

@XXX that went well you lying bastard we all saw through you

.scumbag --> @XXX <-- go get yourself circumcised and bow to mecca 5 times a day, traitor.

Unqualified

Tweets that were abusive and contained no additional information about why the abuse was sent.

@XXX your still a twat though

@XXX spineless weasel.

The results of the coding are shown in the table below.

Table 5: Categories of Abuse

Theme	# Tweets	%
Unqualified	51	34%
Political	37	25%
Personal	33	22%
European Referendum	24	16%
Misclassified	5	3%

Personal abuse and unqualified insults made up over half of abuse levelled at MPs during the collection period (56 percent). Should an 'abuse' algorithm that can automatically detect this kind of language be shown to be effective over an extended duration, there may be an opportunity to use this kind of technology to prevent it reaching politicians.

RESEARCH QUESTION 3: HOW WELL DID THE TECHNOLOGY PERFORM?

In order to better understand the potential of the technology, we evaluated the technological performance of the case studies above.

For technology to be useful, it needs to be capable of presenting accurate, robust information that can be trusted by the user. To test how far that is the case, the technology must be accurate, resilient to change over time, bespoke to the data analysed and transparent. For more information on how the technology developed by Demos and the University of Sussex looks to fulfil this see *Vox Digitas (2014)*.¹

Three technological developments were tested during the case studies: data collection and aggregation, content analysis and demographic analysis.

Data Collection

The data collection methodology showed that it is possible to use open source data to collect valuable information in real time. Even during peaks in activity in late June, the data collection tasks were able to bring all requested tweets into the system. The decision to focus on the tweets sent to and from MPs as the source of data for the study was also felt to be effective, capturing all public activity and all public direct interactions between MPs and other Twitter users.

In the future, however, there are a number of important changes that would be required to successfully develop a tool for MPs. First, data collections on specific issues or topics tailored to the interests of MPs would be necessary. Feedback from MPs emphasised the importance of campaigning on specific issues to their daily work, and that monitoring those conversations would be extremely useful. By adding bespoke keyword-based data collections to the tool, this could be catered to. Second, the collection of data from a wider range of data sources would be vital. By connecting to MPs' public Facebook pages or into Reddit's /Politics/ or /UKPolitics/ boards, the dashboard could provide a one-stop shop into the digital world rather than the fragmented, cross-platform experience currently on offer.

¹ <https://www.demos.co.uk/project/vox-digitas/>

Content Analysis

Two types of content analysis were tested during the pilot: boos and cheers, or 'sentiment' analysis, and a bespoke 'abuse' classifier. Both were successfully deployed when evaluated by their F-scores (see the technical annex) and on the manual sanity check. 'Boos and cheers' was the weaker classifier, with an accuracy of around 66 percent, while the abuse classifier reported an 87 percent accuracy. One explanation for why 'boos and cheers' was less effective is simply how difficult the language of 'sentiment' is; sarcasm and irony are particularly misleading for an algorithm operating without a wider context. Moreover, the algorithm was only as effective as the analysts who trained it, and human analysts will disagree on a significant minority of these decisions.

A cursory glance at the data supports the conclusion that the algorithms worked: politicians who took unpopular decisions, fronted controversial campaigns or were the subject of positive or negative media coverage tended to receive feedback on it through Twitter. Expanding this type of analysis by listening to multiple data sources ought to provide a useful feedback mechanism for politicians in judging how the online public views their actions.

Algorithms are not, however, perfect, and this is especially the case when dealing with subjective subjects like sentiment. This means that when aggregated over a large enough dataset, the overall proportions tend to be accurate, but on a single document the margin for error is greater. On this basis, we would recommend the use of this kind of algorithm in establishing an overall response, rather than a means for identifying individual messages of support or criticism.

Automatically identifying abuse is harder still. The team of analysts made a decision about what they felt was abusive, but this was subjective, and may either underestimate or overestimate what constitutes abuse to an individual MP. This could be counter-productive, as some MPs may feel the tool failed to identify the abuse they were receiving, while others would feel robust or forceful arguments were being unfairly flagged. Again, algorithmic transparency is vital here: it must be clear why a message is flagged or not.

Demographic Analysis

Demographic annotation was a success. Gender annotation was sanity checked and shown to be operating at a high degree of accuracy, and the analysis allowed the user further insight in understanding the dataset. Using the annotator, it was possible to measure, for instance, how male-dominated UK politics is on the platform: 56 percent of the users tweeting in our dataset were classified as male, compared with 33 percent classified as female.^{xiii} This percentage dropped to 30 percent female when measuring total tweets sent – half as many tweets were sent by women during the collection period than by men (59 percent).

Demographic information is particularly important when investigating who an MP is communicating with at any one time, or on any given platform.

Equally important is establishing what differences exist between local, national and international conversations. To do this, we applied a geographic annotator to understand where a user was tweeting from. As noted above, native Twitter geolocation is too sparse to consider using to this end, so algorithmically enriching this data was necessary.

In all, 37 percent of tweets were located to the United Kingdom. 10 percent of these were located at NUTS level 3, roughly comparable to British constituency boundaries.

Although this represents a significant improvement on the numbers geolocated through Twitter's native geotagging (3.5 percent of the total dataset), there is further technology development required to bring the percentage up.

Higher levels of accuracy can be achieved using two forms of off-line analysis: 'network analysis' and 'community analysis'. The principle of network analysis relies on the fact that people (accounts) tend to be associated with other people who share the same demographic profile. People's social networks tend to reflect their physical location, gender, age and interests. It is therefore possible to infer these unknown demographic attributes from the known attributes of other members of your network. Further nuances arise from the fact that there are two types of network: static networks (people who you follow, people who follow you) and dynamic networks (people with whom you exchange messages).

The principle of 'community analysis' arises from the fact that some interactions are tied very closely to a particular geography. Examples of this are interactions with local school, church, political, council or law enforcement organisations. So there are a large number of 'marker' community organisations that help to tie accounts to specific geolocations.

Deploying this additional layer of geographic enrichment could transform the effectiveness of the dashboard in connecting MPs to their constituents.

RECOMMENDATIONS

Technology – in this case big data analytics – is not a panacea for MPs. It can provide a valuable window for MPs to make better sense of new and important digital trends. But it needs to be used carefully and sparingly.

Technological development is only half the story. For digital politics to be a success there will have to be changes in the way politicians and political institutions approach the web, particularly if their aims are to reenergise political participation among the young and disaffected.

Social media and the web will increasingly be the channel through which voters learn about and practice politics. With this comes an expectation that politicians are contactable and responsive online; social media is not just another tool to broadcast a message to passive listeners but a place for conversation and argument. Disregarding this runs the risk of alienating those who treat it in the same way as a letter or an email. Online question and answer sessions, live 'townhall' discussions using new video software such as Periscope or Facebook Live, or crowd-sourcing answers through email or message aggregation show how existing technology can be used to keep up with the demand for two-way communication.

There are risks. Debate and discussion online is frequently chaotic, badly-structured and intimidating, and politicians open themselves up to behaviour they are less likely to experience in the offline world. Alongside technological solutions to improving the level and safety of online debate, politicians need to prepare for the nastier sides of the online world.

Outside of direct communication, the digital world offers an unprecedented opportunity for politicians to listen to the concerns of their constituents. Listening tools along the lines piloted in this study can help politicians identify key issues at both a constituency level and in close-to real time.

For everyday communication, narrowing the focus is also vital. The web provides a good opportunity to be in two places at one time; Westminster and an MP's local constituency. By using automated geographic annotation and prioritisation of constituency-level tweets or messages, contacting local groups with online presences or building online lists of constituents, MPs can cut through the noise of a chaotic national online debate to reach the users they represent and learn about ongoing, everyday issues at a constituency level.

When it comes to questions of national importance, there is an opportunity to make better use of the new digital commons. Feedback and criticism from millions of British web users on existing digital platforms could be aggregated and displayed. Steps should be taken to streamline these platforms to improve their effectiveness: restructuring the petition system, for instance, by consolidating the existing patchwork of sites into a single platform, or aggregating template responses that MPs have complained are frustrating.^{xiii}

Misuse of these platforms by third parties or campaigning organisations is a risk. The extent to which social media networks are influential in voting behaviours is poorly understood, but there have been multiple instances where leading politicians have paid tribute to the power of open data in winning vital votes through identifying swing voters and targeting them surgically, either offline or through targeted advertising. Although much of these companies' power is disputed, their actions risk severe reputational damage to wider digital politics projects.

Beyond this, online political participation is at risk of being frustrated by the inefficiency of existing platforms not tailored to the specific UK political context. We should encourage MPs to experiment with new, dedicated digital democracy technologies. Not only will these technologies provide a better environment for listening and discussion, but they may also allow for political *decisions* to be made. Around the world, governments and local governments have experimented with digital referenda, crowd-sourced policy and 'wikidemocracy', not only to test how far current structures of government might fit into the online world, but to build new structures that make full use of the new opportunities the internet offers.^{xiv}

We therefore conclude that technologists and MPs should work together to build freely available, open source tools for British digital politics.

With that recommendation in mind, we set out below what one such tool might look like (Annex 2 includes some examples of how a front-end dashboard might look). There is existing technology, described above, which could form the basis of a tool that transforms how MPs use and comprehend existing digital channels. Any tool should include the following features:

- *Cross-platform*. Any tool must bring together data feeds from multiple platforms, including Facebook, Twitter and Instagram, providing MPs with a 'one-stop-shop' for social media. Two MPs questioned the utility of a single-platform solution.
- *Customizable*. The tool developed must cater to the interests of a specific MP. Alongside national discussions and mainstream media analysis, each MP will have individual constituency-level discussions to listen to or be involved with and private campaigning interests they will be interested in following. The dashboard must reflect this.
- *Geo-enrichment*. Drawing a distinction between international, national and constituency-level interaction is important. Where that data is not available, the deployment of sophisticated geo-enrichment algorithms is necessary.
- *Bespoke Natural Language Processing*. The core technology tested in the case studies relied on bespoke NLP algorithms, and any further tech development must incorporate this capability. Bespoke classification of free text is the most effective way of allowing users to filter signal from noise and improve their use and experience of social media. Topic identification could be used to group messages by their subject. Urgent messages or questions could be prioritised, while abuse could be flagged. Boos and cheers could be quickly aggregated to provide feedback.
- *Two-way Channel*. The dashboard cannot simply be a listening tool. It must make full use of available APIs to allow the user to post and respond to messages received across platforms.

- *Multi-level.* The dashboard should allow the user to quickly drill down into the data from an overall picture to specific details of interest. For instance, an MP should be able to quickly filter the data based on a certain hashtag or time period.
- *Intuitive and Quick to Use.* Successfully applying visual analytics to a dataset through a dashboard requires the final interface to be comprehensible at a glance. Reports from MPs and their assistants spoke of time pressures in dealing with the online world, and the dashboard must be designed to ease this.
- *Transparent.* What the front-end dashboard does not display is as important as what it does. Decisions around what is visualised must be made publicly. Any algorithms must be periodically tested and their accuracies published to ensure any decisions made on the basis of the dashboard can be trusted.
- *Selective.* Not all data is useful. Identifying the valuable data is a vital prerequisite of building a dashboard and filtering the data made available to the end user. The limitless availability of new data sources means any dashboard is at risk of trying to cover everything. The purpose of the proposed digital dashboard – to allow better listening and communication with voters – should inform and limit the data that is collected.

ANNEX 1: METHODOLOGY

Data Collection

Analysts compiled a database containing all available Twitter handles of the 650 sitting MPs. 562 MPs were found to have an active Twitter account (86 percent). The research team used a technology platform called Method52, developed by CASM technologists based at the Text Analytics Group at the University of Sussex.^{xv} It is designed to allow non-technical researchers to collect data from and analyse very large datasets like Twitter.

Between 9th May – 18th August 2016, we used Twitter's public API to collect all Tweets sent to or from a UK Member of Parliament. In total, this dataset contained 11.4 million tweets at an average of 110,000 a day and from 891,000 unique Twitter accounts. The data was stored on a secure server in JSON format.

Analysis

The dataset was subjected to three types of analysis.

Content Analysis

One of the challenges in analysing social media data is how to categorise and process unstructured text. To achieve this, we used an automated approach involving 'natural language processing' (or NLP). This allows researchers to build models that detect patterns in language use that can be used to undertake meaning-based analysis of large datasets. These were built and applied in different contexts to see where they worked, and where they did not. These models are called 'classifiers'.

Classifiers are built by researchers who train an algorithm to automatically recognise patterns in the text through annotating examples (this is based on linguistic, grammatical, and rules based patterns – not simply word matches). The classifiers then begin to recognise certain patterns and can then automatically spot the same patterns in much larger datasets. NLP is widely used in the analysis of language in 'big data' sets, which are too big for humans to manually analyse, for example, to perform sentiment analysis. The methodology annex includes details of our NLP-based methods.

Classifiers are built by training an algorithm to spot patterns in the language used in the body of the tweet by providing examples. An analyst 'marks up' which category he or she considers a tweet to fall into, and this 'teaches' the algorithm to spot patterns in the language use associated with each category chosen. The algorithm looks for statistical correlations between the language used and the categories assigned to determine the extent to which words and bigrams are indicative of the pre-agreed categories. (For further reading on these methods, see the methodology annex (p.85) in Vox Digitas (2014)).

A number of different classifiers were used in this analysis. A full description of the process and the accuracies of each bespoke classifier is contained in the technical annex.

Boos and Cheers

The boos and cheers classifier – sometimes referred to as 'sentiment' – sought to label tweets as being either broadly supportive of an MP, broadly negative, or neither. The categories were predetermined by analysts and the algorithm was trained on the dataset. Examples of the three categories are shown below, though screennames have been removed.

Cheers

@XXX congrats on this evening's debate. yr comments thruout & closing statement - so passionate & helped me decide to vote leave.

thanks @XXX for supporting our students' exhibition on genocide against the tutsi @houseofcommons today

great session with @XXX mp and our politics students this morning. good luck for your exam on monday @XXX

Boos

@XXX i'm not going anywhere. i was born in this country. but you and your disgusting racist views make me ashamed.

@XXX go back to africa nigger!

it's shameful that having led his country to the brink of economic disaster @XXX has now washed his hands of sorting out the mess

Neutral

@XXX @bbcnews i have been raising this as has @XXX

.@XXX we need to bring forward a law today to allow everyone who couldn't register last night to be allowed to vote in the #euref.

important numbers from @XXX if you're in turkey or have loved ones out there. pls call them if needed

Sentiment analysis is inconsistent across groups of people. A group of ten tweets may be classified differently by two different people. As a result, sentiment analysis is one of the most challenging tasks we can set a computer. To alleviate this, the sentiment algorithm was coded by multiple analysts and disagreements were resolved to ensure consistency across the coders.

The final accuracy of the classifier was 66 percent when categorising tweets as 'boos', 'cheers' or 'neutral'. This allowed analysts to investigate whether there were cues or patterns to the moments users took to Twitter to support or criticise MPs.

Abuse

One case study was the extent to which MPs received abuse through the platform. Identification of abusive tweets required two steps. First, the total dataset was filtered to improve the relevancy of the tweets contained within it. All tweets were annotated by whether the tweet contained an abusive word or words that indicated a judgement. The list of keywords was taken from Google's November 2013 banlist, and both are included in full in the technical annex. This new dataset contained 478,000 tweets.

Second, a classifier was built to determine whether or not the tweet could be classed as abusive. This is an extremely subjective distinction: as with sentiment, a group of humans would not agree on the classification of a group of tweets. As before, the algorithm was coded by multiple analysts and disagreements were resolved to ensure consistency across the coders. Three general principles guided the analysts in determining whether a tweet would be classed as abusive:

First, use of additional expletives ('@XXX you're an ugly cunt! i fucking hate you fat piece of shit.');

second, commands such as 'get out' or

'shut up' ('@XXX get out of st helens, you fraud of a man. '); and third, tweets aimed at 'you', the recipient ('@XXX you are a total idiot. please resign.) when combined with an offensive or derogatory term. Analysts avoided classing disagreement as abusive where it did not meet those thresholds, even when it was forcefully or impolitely expressed.

Examples of tweets classified as abusive are shown below:

*'homosexual' 'married to a jew'???'you cretinous, mouth breathing cunt.
@XXX*

*@XXX my god boy what a complete arse you are. no fucking plan a
never mind plan b! disgraceful behaviour. yellow bastards*

*@XXX fuck of you lying, murdering, cutting, sick disabled unemployed
hating rich loving evil nazi*

The classifier operated at 87 percent accuracy on a two-way split. This allowed analysts to identify abusive messages being sent to MPs and investigate their origins, targets and causes.

Gender

In order to estimate the gender of people posting tweets, we used a pre-existing standard algorithm which is incorporated in Method52. Using a forced-choice approach, it classifies each tweet into one of three categories: 'male', 'female' or 'institution', based on information in the user name and user description fields. When tested in 2015 against a sample of 2,500 users, whose gender was known via traditional survey questioning, this algorithm had an accuracy of approximately 85 percent.

In order to re-test the accuracy of this algorithm on our dataset, an analyst took a random sample of 100 users who had contributed to the dataset, and manually marked them up as 'male', 'female' or 'other/unknown'.

The manual 'sanity check' of this analysis supported these findings. The analyst made their determination based on a review of demographic information, images, tweets and media associated with each account, though there may be a misclassification where a user has deliberately obfuscated their identity in a significant way. The 100 accounts chosen were originally marked by the gender algorithm as above in the

following way. Where a gender could not be determined by an analyst, the user was labelled as 'unknown'.

Tables 6 & 7: Gender Classification (Algorithmic and Manual)

Algorithmically Classified		Manually Classified	
Gender	%	Gender	%
female	44%	female	42%
institution	14%	institution	10%
male	42%	male	45%
		unknown	3%

This suggests that the algorithm performed extremely well on the dataset. For this sample, the algorithm slightly over-estimated the relative proportion of institutional and male accounts relative to female accounts, but the error is small.

Geolocation

To understand where a user was tweeting from it was necessary to geolocate the tweet. If a user has enabled geolocation on their tweets, this information can be extracted from the metadata. However, these users are not common. Of the 3.7 million tweets we collected, just 133,000 had native geotagging enabled (3.5 percent).

To increase the number of users who could be geolocated we enriched the data using a standard geographic annotator found in Method52. The annotator looks for information in other fields (description, location and time zone). When a location is found, it places the tweet into one of a standardised number of geographical regions. These are the 'Nomenclature of Territorial Units for Statistics' (or NUTS). The most general are the 12 NUTS-1 locations that make up the UK, followed by 40 more detailed NUTS-2 locations and 174 NUTS-3 locations. Under tests, it was found to be between 80 percent and 90 percent accurate.

The number of tweets from users that were able to be geographically located to the United Kingdom was increased to 1.4 million (37 percent of the total dataset). 10 percent could be located to NUTS-3 locations.

Metadata Analysis

Metadata is information about a tweet or a user who sent a tweet. Each tweet collected contained this information, such as the time the tweet was sent or whether the tweet was sharing or replying to another Twitter user. A full list of available metadata can be found on dev.twitter.com.^{xvi} This metadata is largely structured, so it depends on numeric analyses, such as establishing an average number of followers or a total number of tweets sent by a user.

Numeric metadata analysis used in the two case studies was performed on follower counts, hashtags, mentions, replies, and retweets.

Follower Counts

Twitter users can receive messages from other Twitter users by 'following' them on the platform. Measuring the number of followers a user has is a useful indicator of the potential reach of the messages they send.

Each user who sent a tweet in our dataset provided information on the number of followers they had at the moment the tweet was sent. This was averaged out across the period for each user. This allowed us to compare the relative potential audiences of individual MPs.

Tweeting using a Hashtag

The hashtags contained in a tweet were also collected and analysed. Use of one or more hashtags on Twitter tends to signify an intent to comment on an issue or join a debate and is a good indicator of the subject of the tweet. During the data collection period, users tweeted using 69,000 different hashtags.

This allowed researchers to compare the frequency with which MPs tweeted, or were tweeted to, about issues or events assigned a hashtag. Examples of this included the hashtags around the EU referendum ([#Brexit](#), [#EURef](#), [#Leave](#) and [#Remain](#)), or those around political television ([#BBCSP](#), [#BBCQT](#)).

Mentioning, Replying to or Retweeting Another User

Mentions, replies and retweets are pieces of Twitter data that indicate an interaction between two users on Twitter. A mention, for instance, is a tweet containing the screenname of another user on Twitter. A reply indicates a tweet by a user has been replied to directly.

When a tweet contained multiple mentions, these were treated separately.

Over the six week period, users mentioned 249,000 users, replied to 103,000 users and retweeted 89,000 users.

Visual Analytics

Building dashboards to interpret data is a type of visual analytics, whereby data too large and complex to understand 'line-by-line' is displayed through visual, often interactive, interfaces. In this study, the hypothesis was that replacing a Twitter feed with appropriate visual interfaces would improve a user's ability to process the information more quickly and intuitively and would reduce the time and effort needed to strip away what is valuable and relevant from that which isn't.

To this end, analysis was considered complete only when the data could be presented through a visual interface. Examples of this are included within each case study. The dashboards were built using Qlikview and Qlik Sense, visual analytics platforms that allowed the researchers to build bespoke interfaces into the data and share them online.

ANNEX 2: THE DASHBOARDS

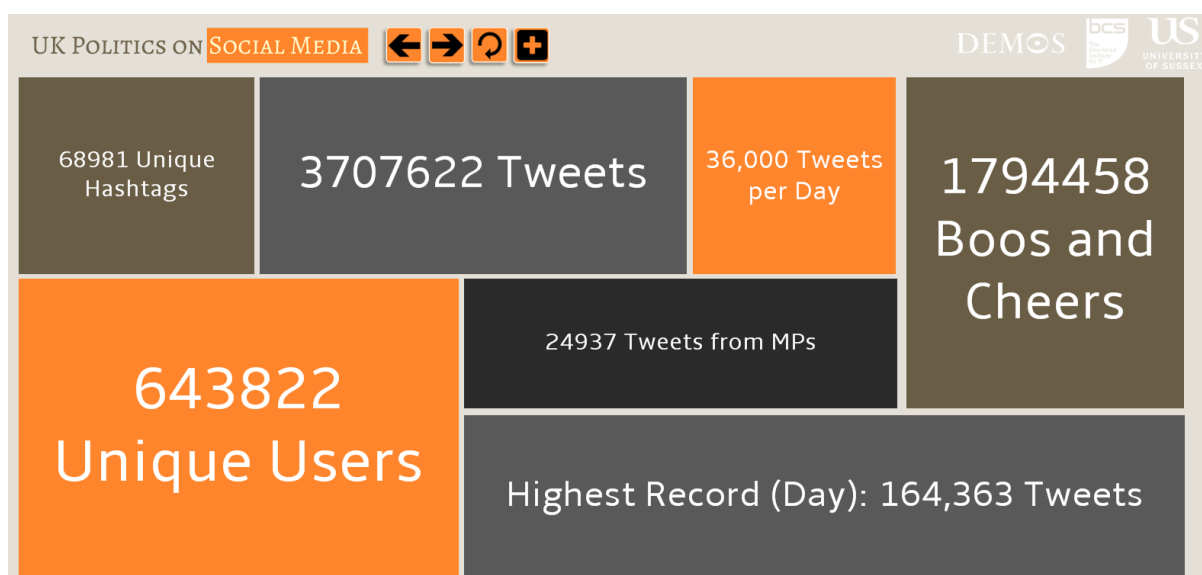
The idea of a digital dashboard for UK politics isn't new. At the time of writing, the government website boasted nearly 900 different dashboards covering everything from stamp duty to cattle tracing.^{xvii} The bespoke iPad app built for the Prime Minister included simple trends from social media and expert commentary alongside other significant statistics.^{xviii} This precedent, alongside a general appetite for this among the MPs we spoke to, suggests that an updated and improved dashboard may be a project worth pursuing.

Nevertheless, we should not lose sight of the risks. A forthcoming Demos report on governance by dashboard warns of the need for "new skills, dynamics, pressures, opportunities and challenges into the practice of

governance”, and the risk in interpreting what's on the screen. “Data presented on dashboard is rarely as straightforward as it appears... [and introduces] a new emphasis on metrics, indicators and measure... [risking] a greater focus on operational issues rather than longer-term strategic ones.”^{xix}

Five dashboards were built during the project as we responded to feedback and explored new ways of using the technology to study the data. Each dashboard was connected to the last, meaning selections made in one impacted the information presented throughout the set.

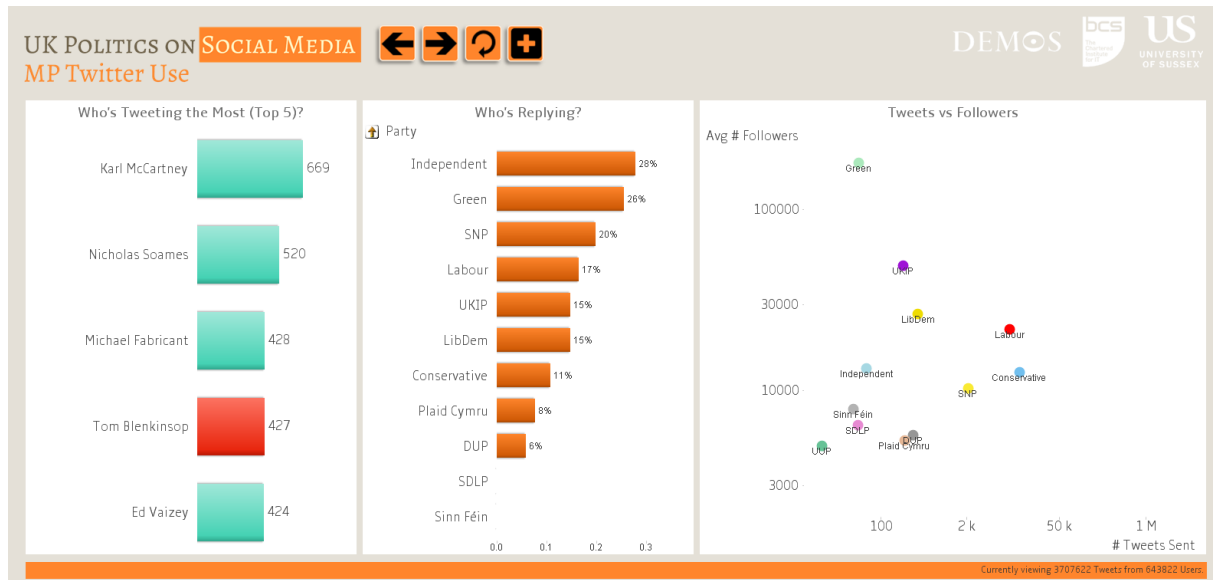
Dashboard 1



The first dashboard looked simply to capture the level of activity across UK politics based on the data we were collecting. This acted as a landing page for a user. This page could be quickly filtered to a certain timeframe or to a certain MP or party to answer basic questions. For instance, a user could identify how many tweets they had received over the last week, the number of users who had contacted them, and whether the tweets were sympathetic or not.

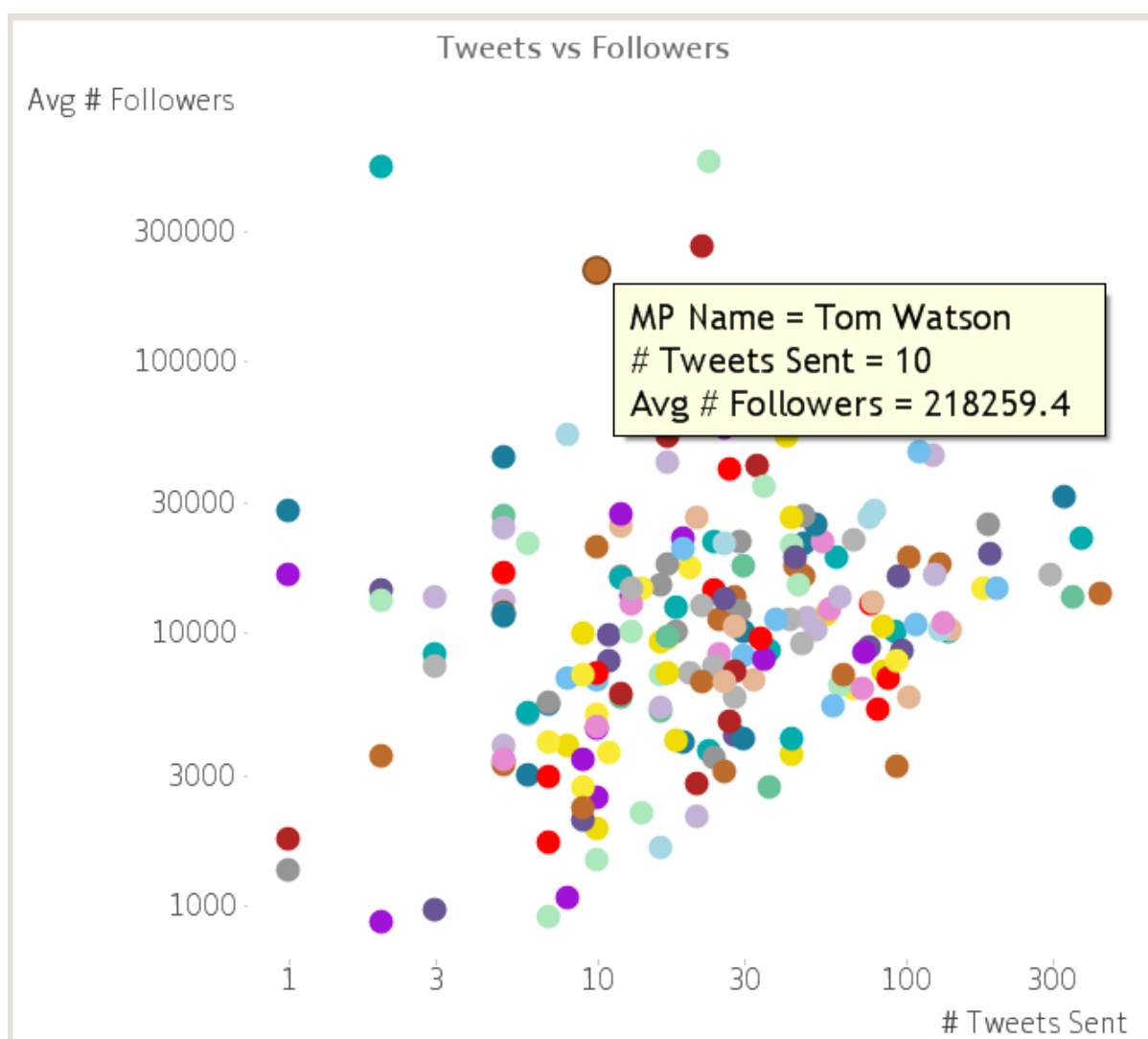
Tiles could also be switched to reflect other data sources or metrics, such as comments from a Facebook page, or the topics of conversation that dominated the period in question. The interface was felt to provide some useful information at a glance, but that the ability to drill down into the data was required.

Dashboard 2



The second dashboard was a window into MP behaviour on Twitter. The visualisations selected were as follows:

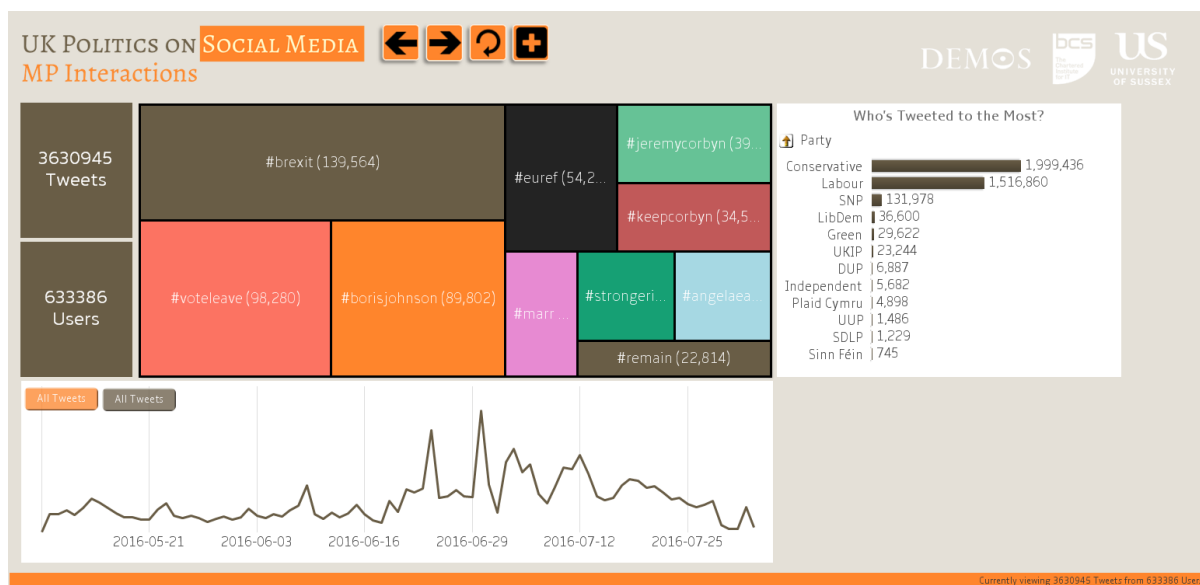
1. A chart that measured the activity of each MP on Twitter, ranking them by the most active MPs and colouring them by party.
2. A chart that measured the proportion of tweets by party that were replies to other Twitter users. Selecting a party would drill into that data. For instance, selecting the Conservative party would show which Conservative MPs were proportionally replying the most.
3. A chart showing the number of tweets sent by each party measured against the average follower count of each MP. Selecting a party would drill into that data, showing the MPs tweeting most frequently and those with the highest followership. This allowed researchers to identify MPs who tweeted frequently but to small audiences, or who had built up large audiences on Twitter but were rarely using the platform. An example of this (using the Labour Party) is shown below.



Colours were used to make the differentiation of each data point easy, but had no analytical significance.

This interface provided the user with new insight. Given an MP, we were quickly able to investigate how their use of the platform compared to their colleagues. We could measure their relative activity, their relative likelihood to use the platform to reply directly to a user or to broadcast, and understand how widely their messages were immediately circulated. All were deemed useful measures in judging an MP's use of a platform, and could be used as the basis for suggestions to improve it.

Dashboard 3



The third dashboard moved the focus of the analysis away from the MPs to the users who were messaging them. Of the 3.7 million tweets captured, 99 percent came from non-MP accounts. The interface was designed to give analysts a window into how Twitter users were interacting with their MPs online. The visualisations chosen were as follows:

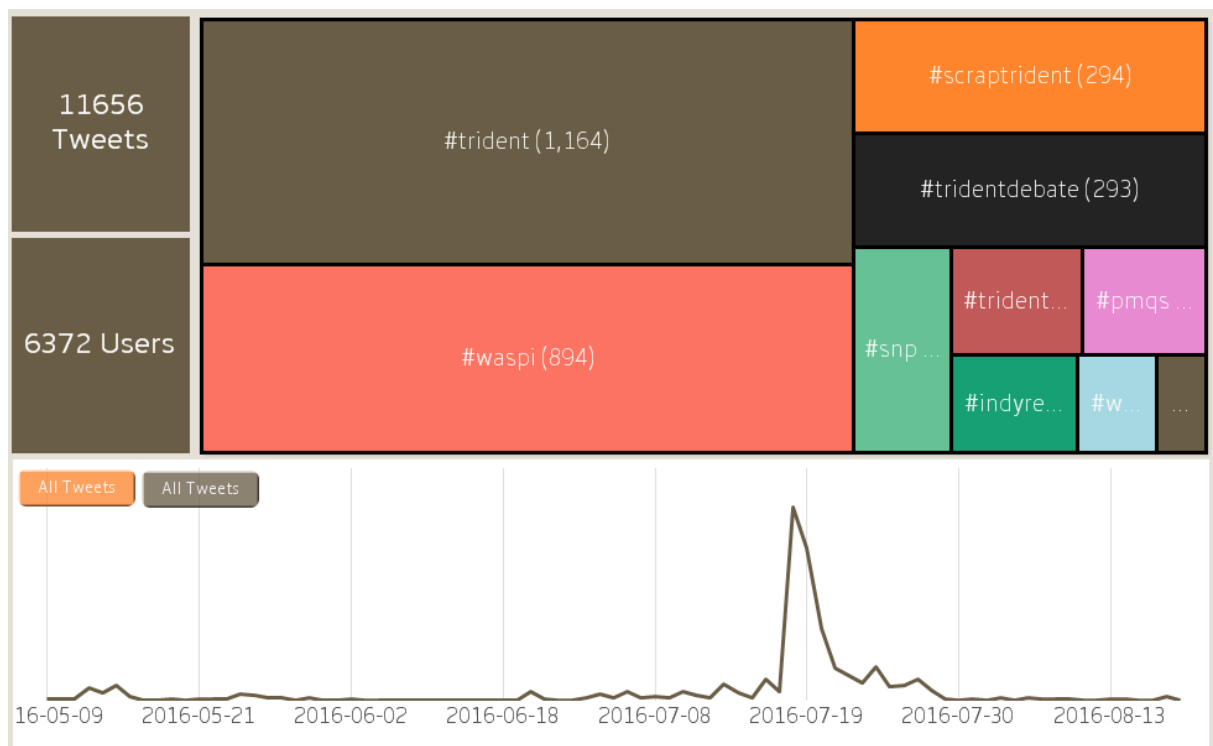
1. Fixed indicators of the number of tweets and users engaging with MPs through the platform.
2. A block chart showing the most widely used hashtags and the number of times each one was used in combination with a tweet directed to or about an MP. Selecting a single hashtag, or a group of hashtags, filtered the rest of the data to show only those tweets.
3. A line chart showing the number of tweets sent each day. Selecting a period on the chart filtered the data to the chosen time period.
4. A bar chart showing which MPs were tweeted to or about the most frequently. This was aggregated at a party level, and selecting a party would drill down into an MP-level view.

This opened up new opportunities to study the data. A peak of Twitter activity, indicated by a spike in the line chart, could be investigated to see why Twitter users had been particularly active during that period. Unsurprisingly, the largest spikes coincided with major political events: the EU referendum results on the 24th June and the subsequent Conservative and Labour leadership elections were very widely discussed.

The subject of the activity can also be inferred from the hashtags used by the Twitter users. Four of the top five most frequently used hashtags were related to the EU referendum.

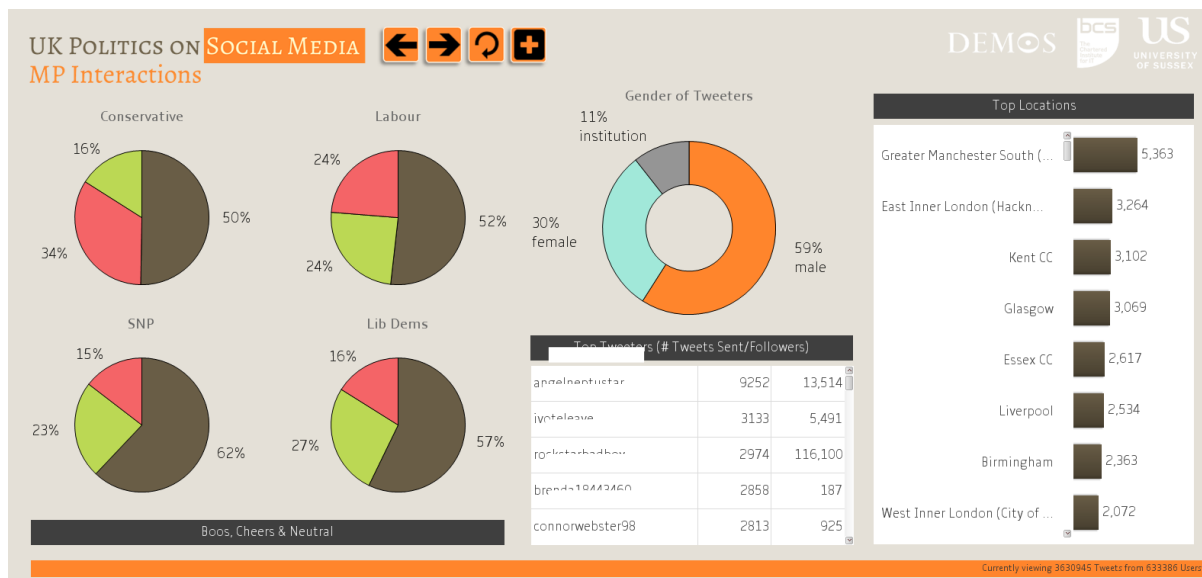
The ability to drill into the data allowed analysts to quickly draw out information around a certain topic, MP, party, or moment in time.

For instance, selecting an MP would show the number of users who had tweeted to or about them, the hashtags used most frequently, and the period in which they received the most attention. An example of this, for the SNP MP Mhairi Black, is shown below.



The hashtags used indicate a different conversation around Mhairi Black's twitter feed by comparison to the overall dataset, with discussions of the Trident nuclear defence program and the Women Against State Pension Inequality (WASPI) debate taking the place of EU Referendum related messaging.

Dashboard 4

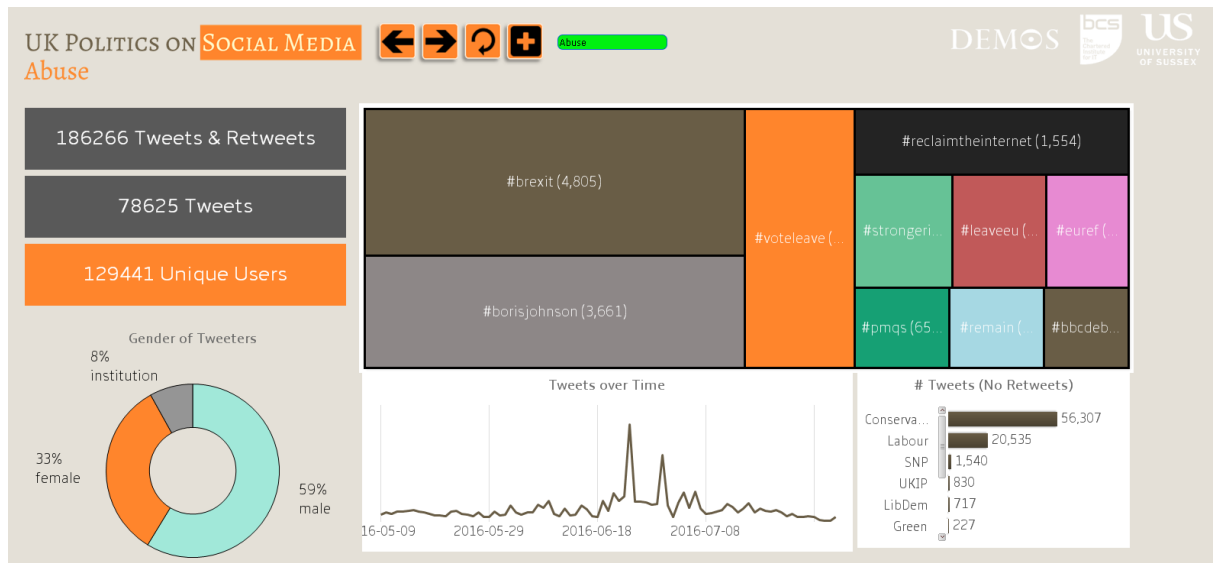


The fourth dashboard added a layer of linguistic analysis, allowing a user to answer specific questions about the data that would require the content of a tweet to be understood and categorised. The selected visualisations were as follows:

1. Four pie charts showing, for the four main UK parties, the proportion of tweets sent to their MPs that were deemed negative ('boos'), positive ('cheers'), or neither ('neutral').
2. A pie chart showing the proportion of users by gender ('male', 'female', or 'institution') as determined by the gender algorithm. Selecting a portion of the chart filters the data to show only tweets from that group.
3. A table showing the users sending the greatest number of messages, and the average number of followers they had.
4. A chart showing the NUTS level 3 locations algorithmically assigned to each user. Selecting a location filtered the data to show only tweets that could be geolocated to that location.

The addition of content analysis allowed the user to compare the ways in which different users responded to MPs. For example, a researcher could quickly compare how favourably female users located in the North of England were messaging Conservative MPs when compared to the overall dataset.

Dashboard 5



The final dashboard was created as an example of how the use of bespoke visualisation software combined with content analysis can be used to investigate a specific issue, in this case the abuse of MPs through Twitter. The visualisations used were taken from previous dashboards. This allowed an analyst to investigate when abuse was taking place, who was receiving it, as well as the topics around which it took place.

This type of dashboard could provide a model for someone monitoring a specific issue (such as abuse of MPs) in real time and respond to the intelligence provided.

ANNEX 3: TECHNOLOGY OVERVIEW

Method52

Data drawn from social media are often too large to fully analyse manually, and also often not amenable to the conventional research methods of social science. The research team used a technology platform called Method52, developed by CASM technologists based at the Text Analytics Group at the University of Sussex.^{xx} It is designed to allow non-technical researchers to analyse very large datasets like Twitter.

Data Analysis

Method52 allows researchers to train algorithms to split apart ('to classify') tweets into categories, according to the meaning of the tweet, and on the basis of the text they contain. To do this, it uses a technology called natural language processing. Natural language processing is a branch of artificial intelligence research, and combines approaches developed in the fields of computer science, applied mathematics, and linguistics.

An analyst 'marks up' which category he or she considers a tweet to fall into, and this 'teaches' the algorithm to spot patterns in the language use associated with each category chosen. The algorithm looks for statistical correlations between the language used and the categories assigned to determine the extent to which words and bigrams are indicative of the pre-defined categories.

The Accuracy of Classifiers

To measure the accuracy of algorithms into the categories chosen by the analyst, we used a 'gold standard' approach. For each, around 100 user descriptions were randomly selected from the relevant dataset to form a gold standard test set for each classifier. These were manually coded into the categories defined above. These tweets were then removed from the main dataset and so were not used to train the classifier.

As the analyst trained the classifier, the software reported back on how accurate the classifier was at categorising the gold standard, as compared to the analyst's decisions. On the basis of this comparison, classifier performance statistics – 'recall', 'precision', and 'F-score' are created and appraised by a human analyst. Each measures the ability

of the classifier to make the same decisions as a human in a different way:

Overall accuracy:

This represents the percentage likelihood of any randomly selected description within the dataset being placed into the appropriate category by the algorithm. It is based on three other measures (below).

Recall:

The number of correct selections that the classifier makes as a proportion of the total correct selections it could have made. If there were 10 relevant descriptions in a dataset, and a relevancy classifier successfully picks eight of them, it has a recall score of 80 per cent.

Precision:

This is the number of correct selections the classifiers makes as a proportion of all the selections it has made. If a relevancy classifier selects 10 descriptions as relevant, and eight of them actually are indeed relevant, it has a precision score of 80 per cent.

F-Score:

All classifiers are a trade-off between recall and precision. Classifiers with a high recall score tend to be less precise, and vice versa. The 'overall' score reconciles precision and recall to create one, overall measurement of performance for each decision branch of the classifier.

The F-score ranges between 0 and 1, with a higher number indicating better performance.

Caveats:

The research of large social media datasets is a reasonably new undertaking. It is important to set out a series of caveats related to the research methodology that the results must be understood in the light of:

- The algorithms used are very good, but not perfect: throughout the report, some of the data will be misclassified. The technology used to analyse tweets is inherently probabilistic, and none of the algorithms trained and used to produce the findings for this paper were 100 percent accurate. The accuracy of all algorithms used in the report are clearly set out in this report.

- Twitter, and especially political Twitter, is not a representative window into British society: Twitter is not evenly used by all parts of British society. It tends to be used by groups that are younger, more socio-economically privileged and more urban. Additionally, the poorest, most marginalised and most vulnerable groups of society are least represented on Twitter.

Tables containing classifier accuracies are shown below.

Abuse Classifier

Label	Precision	Recall	F-Score	Accuracy	Coded
abuse <small>Sample</small>	0.954	0.884	0.918		413
other <small>Sample</small>	0.604	0.806	0.690		246
Unlabelled	0.000	Features	42	0.870	

Boos, Cheers and Neutral

Label	Precision	Recall	F-Score	Accuracy	Coded
positive <small>Sample</small>	0.576	0.559	0.567		478
negative <small>Sample</small>	0.537	0.643	0.585		788
neutral <small>Sample</small>	0.760	0.691	0.724		594
Unlabelled	5987	Features	297	0.655	

Abuse Keyword Annotation List

Threatening or Abusive Language (Analysts' own)

Cut you	Kidnap your	Rape you	Strangle you
Find you	Kill you	Rape your	ur a
Gag you	Kill your	Shoot you	Watching you
Hunt you	Knife you	Shut up	You are
Hurt you	Mutilate	Shut your	You're a
Kidnap	Punch you	Stab your	your a

Offensive Language (Google November 2013 list)

4r5e	ballbag	bollok	clitty litter	cock-sucker	dirty Sanchez	dinks	fuck puppet
50 yard cunt punt	ballsack	boner	clusterfuck	cocksucking	dlck	dirsa	fuck trophy
5h1t	bangbros	boob	cnut	cocksucks	dog-fucker	fanny	fuck yo mama
5hit	bareback	boobs	cock	cocksuka	doggie style	fannyflaps	fuck
a_s_s	bastard	boobbs	cock pocket	cocksukka	doggiestyle	fannyfucker	fucka
a2m	beastial	boooobs	cock snot	cok	doggin	fanyy	fuck-ass
a55	beastiality	boooooobs	cockface	cokmuncher	dogging	fatass	fuck-bitch
amateur	beef curtain	booooooobs	cockhead	coksucka	donkeyribber	fcuk	fucked
anal	bellend	breasts	cockmunch	coon	doosh	fcuker	fucker
anal impaler	bestial	buceta	cockmuncher	cop some wood	duche	fcuking	fuckers
anal leakage	bestiality	bugger	cocks	cornhole	dyke	feck	fuckhead
anilingus	bi+ch	bum	cocksuck	cunts	eat a dick	fecker	fuckheads
anus	biatch	bunny fucker	cocksucked	cuntsicle	eat hair pie	felching	fuckin
ar5e	bimbos	bust a load	corp whore	cunt-struck	ejaculate	fellate	fuckng
arrese	birdlock	busty	cum	cut rope	ejaculated	fellatio	fuckngs
arse	bitch	butt	cum chugger	cyalis	ejaculates	fingerfuck	fuckingshitmotherfucker
arsehole	bitch tit	butt fuck	cum dumpster	cyberfuc	ejaculating	fingerfucked	fuckme
ass	bitcher	butthole	cum freak	cyberfuck	ejaculating	fingerfucker	fuckmeat
ass fuck	bitchers	buttmuch	cum guzzler	cyberfucked	ejaculation	fingerfuckers	fucks
asses	bitches	buttplug	cumdump	cyberfucker	ejakulate	fingerfucking	fucktoy
assfucker	bitchin	c0ck	cummer	cyberfuckers	erotic	fingerfucks	fuckwhit
ass-fucker	bitchng	c0cksucker	cumming	cyberfucking	f u c k	fist fuck	fuckwit
assfukka	bloody	carpet muncher	cums	d1ck	f u c k e r	fistfuck	fudge packer
asshole	blow job	carpetmuncher	cumshot	damn	f_u_c_k	fistfucked	fudgepacker
ass-hole	blow me	cawk	cunilingus	dick	f4nny	fistfucker	fuk
assholes	blow mud	chink	cunilingus	dick hole	facial	fistfuckers	fuker
assmucus	blowjob	choade	cunilingus	dick shy	fag	fistfucking	fukker
assmunch	blowjobs	chota bags	cunt	dickhead	fagging	fistfuckings	fukkin
asswhole	blue waffle	cipa	cunt hair	dildo	faggitt	fistfucks	fuks
autoerotic	blumpkin	d1t	cuntbag	dildos	faggot	flange	fukwhit
b1tch	boiolas	clit	cuntlick	dink	faggs	flog the log	fukwit
b00bs	bollock	dit licker	cuntlicker	gangbang	fagot	fook	fuX
b17ch	cocksucker	clitoris	pornos	pissers	fagots	fooker	fuX0r
b1tch	cuntlicking	clits	prick	pisses	fags	fuck hole	gangbang
gang-bang	jiz	ma45terbate	motherfucking	nobjokey	pricks	phuq	spac
gangbanged	jizm	ma5terb8	motherfuckings	numbnuts	pron	pigfucker	spunk
gangbangs	jizz	ma5terbate	motherfuckka	nut butter	pube	shagger	t1tt1e5
gassy ass	kawk	mafugly	motherfucks	nutsack	pusse	shaggin	t1tties
gaylord	kinky Jesus	masochist	muff	omg	pussi	shagging	teets
gaysex	knob	masterb8	muff puff	orgasim	pussies	shemale	teez
goatse	knob end	masterbat*	mutha	orgasims	pussy	shi+	testical
god	knobead	masterbat3	muthafecker	orgasm	pussy fart	shit	testicle
god damn	knobed	masterbate	muthafucker	orgasms	pussy palace	shit fucker	tit
god-dam	knobend	master-bate	muther	p0rn	pussys	shitdick	tit wank
goddamn	knobhead	masterbation	mutherfucker	pawn	queaf	shite	titfuck
goddamned	knobjocky	masterbations	n1gga	pecker	queer	shited	tits
god-damned	knobjokey	masturbate	n1gger	penis	rectum	shitey	titt
ham flap	kock	mof0	nazi	penisfucker	retard	shitfuck	tittie5
hardcoresex	kondum	mof0	need the dick	phonesex	rimjaw	shitfull	tittiefucker
hell	kondums	mo-fo	nigg3r	phuck	rimming	shithead	titties
heshe	kum	mothafuck	nigg4h	phuk	s hit	shiting	tittyfuck
hoar	kummer	mothafucka	nigga	phuked	s.o.b.	shittings	tittywank
hoare	kumming	mothafuckas	niggah	phuking	s_h_i_t	shits	titwank
hoer	kums	mothafuckaz	niggas	phukked	sadism	shitted	tosser
homo	kunilingus	mothafucked	niggaz	phukking	sadist	shitter	turd
homoerotic	kwif	mothafucker	nigger	phuks	sandbar	shitters	tw4t
hore	l3i+ch	mothafuckers	niggers	pimpis	sausage queen	shitting	twat
horniest	l3itc	mothafuckin	nob	piss	schlong	shittings	twathead
horny	labia	mothafucking	nob jokey	wank	screwing	shitty	twatty
hotsex	LEN	mothafuckings	nobhead	wanker	scroat	skank	twunt
how to kill	lmao	mothafucks	nobjocky	wanky	scrote	slope	twunter
how to murder	lmfao	mother fucker	pisser	whoar	scrotum	slut	v14gra
jackoff	lust	mother fucker	pissflaps	whore	semen	slut bucket	v1gra
jack-off	lusting	motherfuck	pissin	willies	sex	sluts	vagina
jap	mOf0	motherfucked	pissing	willy	sh!+	smegma	viagra
jerk	mOf0	motherfucker	pissoff	wtf	sh!t	smut	vulva
jerk-off	porn	motherfuckers	poop	xrated	sh1t	snatch	w00se
jism	porno	motherfuckin	pornography	xxx	shag	son-of-a-bitch	wang

Notes

1

<https://www.ons.gov.uk/businessindustryandtrade/itandinternetindustry/bulletins/internetusers/2016>

¹ <https://www.statista.com/statistics/271348/facebook-users-in-the-united-kingdom-uk-by-age/>

² https://www.ofcom.org.uk/data/assets/pdf_file/0014/82112/2015_adults_media_use_and_attitudes_report.pdf

³ <https://www.bdo.co.uk/en-gb/insights/industries/public-sector/a-review-of-social-media-usage-in-uk-local-government>

⁴ <http://www.digitaldemocracy.parliament.uk/documents/Open-Up-Digital-Democracy-Report.pdf>

⁵ Carl Miller, The Rise of Digital Politics (2016) <https://www.demos.co.uk/wp-content/uploads/2016/10/Demos-Rise-of-Digital-Politics.pdf>

⁶ <https://www.facebook.com/OfficialBritainFirst/>

⁷

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/492972/gs-16-1-distributed-ledger-technology.pdf

⁸ <https://www.ipsos-mori.com/Assets/Docs/Polls/ipsos-mori-veracity-index-2015-tables.pdf>

⁹ Talking to ourselves (Krasodonski-Jones, 2017)

¹⁰ Recent changes in how society considers gender to be 'non-binary' may require a rethink in how gender analytics is performed.

¹¹ <https://www.ft.com/content/499e3ac8-c50e-11e5-808f-8231cd71622e>

¹² http://paginas.fe.up.pt/~feliz/e_poster4_wiki-law-government.pdf

¹³ This group is led by Professor David Weir and Dr Jeremy Reffin. More information is available about their work at: www.taglaboratory.org

¹⁴ <https://dev.twitter.com/overview/api/tweets>

¹⁵ <https://www.gov.uk/performance>

¹⁶ <https://www.gov.uk/government/publications/case-study-on-action-4-digital-capability-across-departments/action-4-case-study-digital-capability-across-departments--2>

¹⁷ Governance by Dashboard, Demos (2017)

¹⁸ This group is led by Professor David Weir and Dr Jeremy Reffin. More information is available about their work at: www.taglaboratory.org

The Centre for the Analysis of Social Media is a collaboration between Demos and the University of Sussex. The Centre combines automated data extraction and sentiment analysis with social science statistics, analysis and ethics, to produce insightful and robust policy research.

Alex is a researcher of the Centre for the Analysis of Social Media. His primary research interest is political extremism and its reportage on social media. He also manages CASM's analytics capability, including data collection, analytics and visualisation.