Department for Work & Pensions

GSR GOVERNMENT SOCIAL RESEARCH

Social Science in Government

The Use of Social Media for Research and Analysis: A Feasibility Study

December 2014

DWP ad hoc research report no. 13

A report of research carried out by the Oxford Internet Institute on behalf of the Department for Work and Pensions

This document/publication is also available on our website at:

https://www.gov.uk/government/organisations/department-for-work-pensions/about/research#research-publications

If you would like to know more about DWP research, please email:
Socialresearch@dwp.gsi.gov.uk

First published 2014.

ISBN 978-1-78425-407-0

Summary

- The aim of this report is to explore the ways in which data generated by social media platforms can be used to support social research and analysis at the Department for Work and Pensions [DWP]. The report combines a general review of all the possibilities generated by social media data with an empirical exploration assessing the feasibility of some solutions, focussing in particular on the examples of Universal Credit and Personal Independence Payment.
- The report argues that social media data can be useful for social research purposes in two key respects. Firstly, these media can provide indications of information seeking behaviour (which may indicate public awareness of and attention to specific policies, as well as providing an idea of the sources where they get information from). Secondly, they can provide indications of public opinion of specific policies, or reaction to specific media events. This means that social media are positioned to provide social researchers at the DWP with a variety of useful data sources, such as:
    - Indications of public reaction to specific policy announcements or proposals
    - Insight into public experiences of services the DWP provide (such as Jobcentres)
    - Ways of measuring overall public attention to DWP policies, or awareness of key policy changes
    - Ways of measuring general social trends of importance to the DWP, such as a rise in employment
    - Insight into the sources of public opinion, such as where people get information on specific DWP policies
- However we also argue that caution is needed in interpreting the results of social media data, or generalizing from these data to the public at large. The science behind many of these methods is still developing; and major questions remain how to employ them properly.
- Overall, we recommend that all social media data be “benchmarked” against other data sources as most indicators developed in the report are very difficult to interpret in isolation.
- This report does not focus on using social media for public relations or communications, rather it looks at its role in social research and analysis.

Contents

List of tables and figures

Acknowledgements

The report's authors would like to acknowledge the financial support provided by the Department for Work and Pensions which commissioned the report, and the wide range of useful critiques and comments provided by Andrea Kirkpatrick, Dimitra Karperou, and several anonymous reviewers.

We would also like to thank Wikimedia Deutschland e.V. and the Wikimedia Foundation which provided access to the Wikipedia data used in this report; and i2-sm whose data collection platform was used to access Twitter and Facebook data.

Finally, we would like to thank the technical staff at the Oxford Internet Institute for infrastructural support to the project.

All errors and omissions remain our own.

The Authors

Jonathan Bright is a Research Fellow at the Oxford Internet Institute, University of Oxford. He is a political scientist specialising in political communication and computational social science (especially “big data” research”). His major research interest is around how people get information about politics, and how this process is changing in the internet era.

Helen Margetts is Director of the Oxford Internet Institute, University of Oxford, where she is Professor of Society and the Internet, and a Fellow of Mansfield College. She is a political scientist specialising in digital government and internet-mediated collective action. She is co-author (with Patrick Dunleavy) of Digital Era Governance: IT Corporations, the State and e-Government (Oxford University Press, 2006, 2008).

Scott Hale is a Research Assistant and doctoral candidate at the Oxford Internet Institute, University of Oxford. He is interested in developing and applying computer science methods to social science questions in the areas of political science, language, and network analysis.

Taha Yasseri is a Researcher at the Oxford Internet Institute, University of Oxford. His main research focus is on online societies, government-citizen interactions on the web and structural evolution of the web. He uses mathematical models and data analysis to study social systems quantitatively.

Glossary and Abbreviations

Abbreviations

- DWP – The Department for Work and Pensions
- API – Application Programming Interface – An interface for accessing the data contained in a web platform or service in a systematic way. Many social media sites maintain such an API.
- PIP – Personal Independence Payments
- JSA – Jobseeker’s Allowance
- DLA – Disability Living Allowance
- TOS – Terms of Service

Glossary

- Sentiment analysis – A family of techniques which aim to automatically extract “sentiment” from a piece of text (for example, whether it is positive or negative, or whether the person writing it was angry or excited, etc.).
- Social media – A means of communication, based around a website or internet service, where the content being communicated is produced by the people using the service. Can be distinguished from other types of media (such as the news media) where there is a clear distinction between the producers and consumers of content.
- Social network – A type of social media which, in addition to the features listed above, allows users to maintain lists of other individuals on the site with which they want to maintain special or frequent contact

Summary

The aim of this report is to explore the ways in which data generated by social media platforms can be used to support social research and analysis at the Department for Work and Pensions [DWP]. The report combines a general review of all the possibilities generated by social media data with an empirical exploration assessing the feasibility of some solutions, focussing in particular on the examples of Universal Credit and Personal Independence Payment.

The report argues that social media data can be useful for social research purposes in two key respects. Firstly, these media can provide indications of information seeking behaviour (which may indicate public awareness of and attention to specific policies, as well as providing an idea of the sources where they get information from). Secondly, they can provide indications of public opinion of specific policies, or reaction to specific media events. This means that social media are positioned to provide social researchers at the DWP with a variety of useful data sources, such as:

- Indications of public reaction to specific policy announcements or proposals
- Insight into public experiences of services the DWP provide (such as Jobcentres)
- Ways of measuring overall public attention to DWP policies, or awareness of key policy changes
- Ways of measuring general social trends of importance to the DWP, such as a rise in employment
- Insight into the sources of public opinion, such as where people get information on specific DWP policies.

Development of methodology in these areas remains in its early stages, with a focus on experimentation. One of the key aims of the report was to point out specific areas where the DWP itself could start to experiment. In particular, we suggest that the DWP looks at the following areas:

- Google Trends data provides useful indicators of how many people are thinking about a topic at any given time. This is a useful way of measuring the extent to which the public is aware of new policies (for example, Universal Credit), and may also give short term indications of upcoming changes in, for example, the number of jobseekers.
- Google Search data provides useful indicators of where the public gets information on particular DWP policies. This could allow the DWP to fine tune its communication strategy, and also tell them which other

websites are currently informing the public on DWP policies (including, for example, private sector sites over which DWP has no control).

- As well as Google, Wikipedia provides a further rich source of information on how many people are interested in a given policy or initiative.
- General social media platforms such as Twitter and Facebook provide a means of assessing how many people are discussing any policy at a given time, and also a way of potentially evaluating their sentiment towards that policy. These platforms can provide a useful indication of the impact of specific media events or press releases.
- The DWP's own social media accounts, run by their network of local Jobcentres, provide a further resource in this regard. They may be useful both for recording feedback and contacting specific regional populations. However, the overall level of activity around these accounts varies, meaning care should be taken when it comes to evaluating specific policy proposals.
- We also highlight the importance of "themed" social network sites such as Mumsnet, though these sites did not fall in the scope of the empirical section of the report.

Some of the above can also be analysed using more traditional techniques such as survey analysis. In comparison with such surveys, social media data present the following advantages.

- They are comparatively cheap to collect, when compared with the cost of traditional sample surveying.
- They can be collected and analysed quickly, once systems are put in place which perform such analysis.
- They offer potentially very large samples, which means that questions about smaller "subgroups" can be accessed (for example issues affecting a certain geographical area).

However we also argue that caution is needed in interpreting the results of social media data, or generalising from these data to the public at large. Social media sources do not immediately replace other sources of social research data such as sample surveys: the science behind many of these methods is still developing; and major questions remain around how to employ them properly. In particular, the following challenges are important:

- Social media users are not representative of the population at large (their use is much more widespread amongst the young, for instance).
- Different people use these media in different ways: most of the postings to social media are made by a subset of overall users, who again are likely to suffer from problems of representativeness.

- Techniques for the automatic extraction of opinions from social media are still developing and have to be interpreted with caution.

Overall, we argue that as social media are increasingly embedded in the fabric of life, it is increasingly difficult to ignore the potential they present for social research which can inform policy-making and service delivery, providing data in both quantities and richness that would be prohibitively expensive to duplicate with traditional survey research. However, we also argue that research is needed before social media data can be integrated into DWP working practices. In this context in particular, we recommend that all social media data be "benchmarked", as most indicators developed in the report are very difficult to interpret in isolation and require on-going data collection. This is a process which may involve:

- Comparing social media indicators over time (for example, if positive or negative sentiment is increasing or decreasing).
- Comparing social media indicators across Departments or policies (for example, seeing if one Jobcentre is attracting more social media attention than another).
- Linking these data with other sources of socially generated information, for example web traffic to DWP's own information and jobs websites.
- Linking these data with other trusted sources of information generated with more traditional techniques (for example survey data, ONS employment statistics).

1 Introduction

The main aim of this report, and the research project on which it is based, is to explore how social media data could be used to help support policy research and analysis at the Department for Work and Pensions [DWP]. This general question is also allied to a more specific objective, which is to explore the extent to which such data could act as a reliable source of public opinion in respect of two key Welfare Reform policies: Universal Credit and Personal Independence Payment.

In this report, we adopt a broad definition of social media, as simply any site where the content is created by the users of the website, and hence cover a wide variety of different types of social media in our review, from chatrooms and discussion forums to search engines and microblogging outlets. However we also recognise that contemporary interest in social media in the academic, public and business sectors has been driven largely by the rise of a limited number of “mass” social media platforms, such as Twitter, Facebook and YouTube, because of the astonishing penetration rates which they have achieved and their relatively open stance towards distributing data. The opportunities created by this smaller number of platforms hence constitute the bulk of the report. We also consider the wider topic of “socially generated data”, which may emerge on platforms such as Google which are not typically considered social media but may nevertheless offer useful insights into the opinions and behaviours of the general public. As these types of social media start to play an ever greater role in our lives, the breadth and richness of the data they provide on human behaviour is increasing rapidly. A wide variety of private organisations are starting to exploit these data for research purposes, and we argue that governmental bodies should be doing the same.

This report is structured in the following way. Part 1 offers a general overview of social media. It begins by defining what is meant by the term, and describing what is known about who uses social media and how penetration rates are changing over time. It then moves on to look at some of the opportunities social media data has created for research in both the business and academic sectors, looking in particular at the use of social media as a measure of public attention to particular issues, and the use of social media as a source of public opinion on a range of questions. Here we also highlight key methodological challenges involved in social media research, and potential strategies for overcoming them. Finally, this section reviews technical, legal and ethical challenges inherent in the use of social media data.

Part 2 looks more specifically at potential uses of social media for the DWP. Building on part 1, we look at how techniques for measuring public attention and public opinion could potentially be deployed to support DWP work. Over four sections, we look at the potential applications of data from Google, Wikipedia, Facebook and Twitter for shedding light on a range of questions in which the DWP may be interested. We focus our examples on the cases of Universal Credit and Personal Independence Payment as they are two currently live policy projects, though it is our intention that our suggestions could be generalised to other policy areas.

We conclude by setting out our recommendations for the DWP's future engagement with social media for research purposes (though we should make it clear at the outset that this report is not based on an in-depth consideration of current DWP organisational structure or working practices - rather, it is a feasibility study designed to highlight in a general way the potential of social media analytics for research purposes). Social media research is in its early stages, and many methodological challenges remain in terms of properly integrating the data provided by such platforms into DWP work. We caution against the simple interpretation of social media metrics for social research such as "the mood on Twitter", which may be quite different to the mood of the public as a whole. Nevertheless the possibilities presented by social media are considerable, providing data in both quantities and richness that would be prohibitively expensive to duplicate with traditional survey research. Overall we argue that as social media are increasingly embedded in the fabric of life, it is increasingly difficult to ignore the potential they present for social research which can inform policy-making and service delivery.

2 Part 1 - Social Media for Social Research

In this first section we provide a general overview of the use of social media data for social research. We begin by offering a definition of social media, which we keep purposefully broad, including not only “mass” social media platforms such as Twitter and Facebook, but also more narrow social forums (such as Mumsnet, a website for parents), and sites which, while not really social *media*, nevertheless hold large amounts of “socially generated data” (such as Google). We then move on to describing how the data generated by these platforms can be useful for social research. Though applications are still in their early stages, we highlight the potential of social media to provide an indication of what topics the public is currently thinking about, where they look for information about these topics, and what opinion they may have on them. Many of these questions can also be explored through more traditional social research methods such as the sample survey; hence in this section we also compare the use of social media data with survey methods, highlighting both potential strengths and some weaknesses.

2.1 Defining Social Media

If media are simply means of communication, “social” media may be defined as websites or other internet based services where the content being communicated is created by the people who use the service. Unlike, for example, a news website, where the content is created by a journalistic and editorial staff for mass consumption, on social media sites there is no clear cut separation between producer and consumer.[1] Different social media sites structure these production and consumption roles differently. Some sites, such as Wikipedia, award status based on the amount previously produced for the community, and only allow users with a certain status to take certain types of action. Other sites such as Facebook allow essentially everyone to create the same kind of content. Furthermore, regardless of the presence of specific constraints, users themselves tend to structure their interaction with social media sites in different ways, with a handful of users contributing extensively to the sites, whilst the majority contribute rarely or never.

[1] Bruns, A. (2008.) The Active Audience: Transforming Journalism from Gatekeeping to Gatewatching. In *Making Online News: The Ethnography of New Media Production*. Eds. Chris Paterson and David Domingo. New York: Peter Lang

Within this broad definition of social media, a variety of more specific sites can be found. Two major factors are of importance in defining the extent to which sites are different. Firstly, there is the way in which the site in question manages the identities of its users as a way of both enabling and constraining access to content. Some sites, often ones which focus on information provision such as Wikipedia, do not require the creation of an account to access material (though accounts are encouraged to post content). Others such as Facebook require an account before most data can be accessed. Furthermore, many account based sites also encourage users to create connections between other users on the site; connections which can be a means of filtering content. In this type of site, which have sometimes been more narrowly defined as "social networking sites",[2] content is created by the entire community, but the actual content a user sees is created by people which that user has chosen to connect with. Different sites also structure link creation in different ways: for example, on Twitter, users can choose to receive content from anyone, whereas on Facebook they must reciprocally confirm each other as "friends" before many types of content can be viewed.

Secondly, some social media websites dedicate themselves to a specific theme or niche interest, whilst others attempt to create a more general type of space for social interaction (within which more specific niches can spring up). Social networking sites such as Facebook and Twitter, for example, are generalist: a wide variety of social interactions can take place on them. Other sites such as LinkedIn (a site designed for professional connections) or Mumsnet (a site designed for parents to meet and discuss) have more of a specific theme. Here, though the community is open to anyone, a certain type of content creation is encouraged; and moderators (again perhaps drawn from the community) may actively shut out people who are posting things which are considered off topic.

Finally, in this context we should also mention sites which fall slightly outside what is traditionally conceived of as social *media*, as their primary purpose is not to facilitate communication, but which nevertheless store large amounts of "socially generated data". An obvious example here would be Google: by recording the search interests of (literally) billions of people, there is an important sense in which Google is a "socially" generated site.

2.1.1 Social Media Usage

On the basis of the above definition, many types of technology which have been supported by the internet (and indeed non-internet technologies) could be conceived of as social media. Certainly social media have been part of the internet since its earliest foundational days, in the form of things such

[2] boyd, danahd and Ellison, N. (2007). Social Network Sites: Definition, History and Scholarship. *Journal of Computer-Mediated Communication*, 13(1), 1, 210-230

as usenet discussion groups, whilst the World Wide Web itself was also conceived of in broadly social terms when it was created.[3] Hence there is a sense in which social media have a long history.

However current academic and industry interest in social media is much more recent (dating back around five years). This interest has been driven by the rapidly broadening user base for social media technologies, which is of course related to the continuing spread of internet use itself (approximately 73% of the UK's population accessed the internet every day in 2013[4]). The rise in social media use has been rapid: in 2011, approximately 60% of internet users were also social media users, up from just 17% in 2007.[5] Much of this change has been driven by the emergence of a small number of "mass appeal" social media websites, of which Twitter and Facebook are the obvious examples. These sites are characterised by their ease of use, their generic nature (i.e. they eschew focus on a particular subject or area of interest) and their wide penetration, meaning that significant portions of the population have created an account (though it should be remembered that not all of these people use these accounts regularly). This mass usage creates a kind of "network effect" which helps promote further uptake, as people now have an incentive to join Facebook because many of their friends are already there.[6] A March 2013 survey of the size of Facebook, which is by far the biggest social network in terms of user numbers, estimated the number of UK accounts at over 30 million (though it should be noted that an account does not necessarily equal a "person": some people may operate multiple accounts, businesses and organisations also run accounts and for various reasons others create fake accounts).

Of course, there are no guarantees that sites such as Twitter and Facebook will continue to grow and prosper. Recent history presents several examples of sites, such as Bebo and Myspace, which went from enjoying significant usage bases to relative obscurity in a short space of time (in 2007, for example, Myspace received more visits than Google did in the US). However there is also a sense in which the current leaders appear to have learnt from some of the mistakes made by these previously dominant sites, and certainly appear to be well positioned to last for a significant amount of time.

It is important to note, especially when considering the use of social media data for research purposes, that social media usage is not distributed evenly across the population. It is particularly imbalanced across different age

---

[3] See http://first-website.web.cern.ch/

[4] See Office of National Statistics. (2013). "Internet Access – Households and Individuals, 2013", Available from: http://www.ons.gov.uk/ons/dcp171778_322713.pdf

[5] See Dutton, W., and Blank, G. (2011). "Oxford Internet Survey 2011: Next Generation Internet Users", 34. Available from: http://oxis.oii.ox.ac.uk/sites/oxis.oii.ox.ac.uk/files/content/files/publications/oxis2011_report.pdf

[6] See Dutton, W., and Blank, G. (2011), 33

groups.[7] Social networking site use is very common amongst younger age groups: over 80% of internet users between 14-34 are also social network users, while just 20% of internet users over 65 are.[8] Other typical demographic characteristics show less of an imbalance. Some studies have shown that different gender groups use different social media sites differently: with, for example, men more likely to use the business oriented network LinkedIn and women more likely to use visual sharing tool Pinterest.[9] However recent surveys in the UK context have shown no difference in usage rates across gender overall.[10] Of particular significance for the DWP, early findings from the 2013 Oxford Internet Survey suggest that the biggest increases in internet use were seen in low-income households (58% of households earning less than £12,000 a year use the Internet, up from 43% in 2011) and there did not seem to be significant difference in social media usage between employed and unemployed internet users.[11]

Perhaps more important than the demographics of overall access to social media sites are the diverging uses towards which these sites are put. In particular, many studies have reported diverging patterns in terms of who actively creates online content, how often they create it, and what specific types of content they create. Across many different social media sites, there is a consistent pattern whereby the majority of content is created by a relatively small minority of users[12]. 50% of social media users only visit the sites on a monthly basis[13]. Sometimes this difference has been conceptualised as the difference between “residents”, who actively spend their lives online, and “visitors”, who go online to search out particular bits of information, and leave when they have found it.[14] Research in this area is in its early stages so it would be inappropriate to draw firm conclusions about why some people are more likely to engage in content creation than others, especially in the context of a technological environment which is still changing rapidly. Nevertheless, early results indicate that age, education and “skills” appear to have some

[7] Unless otherwise stated, all further statistics in this section refer to internet users, rather than the whole population
[8] See Dutton, W., and Blank, G. (2011), 36
[9] Acquisti, A. and Gross, R. 2006. "Imagined communities: Awareness, information sharing, and privacy on the Facebook". *Privacy Enhancing Technologies, Lecture Notes in Computer Science*, 4258, 36–58
[10] Office of National Statistics. (2012). “Internet Access - Households and Individuals, 2012 part 2.". Available from: http://www.ons.gov.uk/ons/rel/rdit2/internet-access---households-and-individuals/2012-part-2/stb-ia-2012part2.html
[11] Dutton, W., and Blank, G. (2013). “OxIS 2013 Report: Cultures of the Internet”. Available from:
http://oxis.oii.ox.ac.uk/sites/oxis.oii.ox.ac.uk/files/content/files/publications/OxIS_2013.pdf
[12] See, for example, Ortega, F., Gonzalez-Barahona, J., and Robles, G. (2008). "On the inequality of contributions to Wikipedia.". *HICSS '08 Proceedings of the 41st Annual Hawaii International Conference on System Sciences*., 304
[13] Dutton, W., and Blank, G. (2011), 35
[14] White, D., and Le Cornu, A. (2011). "Visitors and Residents: A new typology for online engagement". *First Monday*, 16(9).

kind of influence on content creation.[15] These effects change when considering different types of content: those with higher academic qualifications are more likely to create political content or engage in political discussion, and less likely to create social or entertainment type content.[16] Furthermore, a variety of psychological characteristics have been found to be associated with high levels of social media use, with people considered “extroverts” in particular more likely to create content about themselves. Finally, it is also worth noting that usage patterns of social media are not uniform across time.[17] Though they are often a location for internet access in general, many workplaces discourage or even outright block the use of social media during working hours: hence many people will connect only on evenings or weekends (though the emergence of mobile devices as a means for accessing social media is starting to change this).

Hence while social media as a whole count on usage from large sections of the population, a random sample of social media users will not be representative of the population at large in many respects (though certainly more reflective than it would have been in the 1990s). The issues this may create for using social media data for research are discussed further below.

2.2 Using Social Media for Social Research

Currently, the major practical use of social media for both business and governments is as a means of managing public relations. Social media provide a channel where organisations can quickly diffuse particular messages of interest to a wide audience (compared with other diffusion channels such as press releases or paid for advertising). They also constitute an arena where the issues of the day are frequently debated and where opinions can be formed on a wide range of topics. Hence many large organisations now have social media teams in their communications or public relations departments which both monitor current events on social networks and actively release content to those networks.

This report, however, does not aim to address the use of social media for public relations. Rather, we focus on a secondary usage, which is still developing but has equal if not greater potential: the use of social media data for social research, particularly as an alternative to public surveys and opinion polls. In this section, we explore two main applications of social media data for social research: as a way of knowing what the public are currently thinking

[15] Hargittai, E., and Walejko, G. (2008). "The Participation Divide". *Information, Communication and Society*, 11(2), 239-256.
[16] Blank, G. (2013). "Who creates content?" *Information, Communication and Society*, 11(2), 590-612.
[17] Yasseri, T., Sumi, R., and Kertész, J. (2012). "Circadian Patterns of Wikipedia Editorial Activity: A Demographic Analysis.". *PLoS One*, 7(1).

about (and hence perhaps predicting what they are about to do); and as a way of analysing public opinion or sentiment on specific issues. As the mass uptake of social media described above is relatively new, these uses themselves are still developing, with many more applications likely to emerge. However these two categories cover in many ways the current state of the art.

Of course, social media data are not the only ways to explore such questions: many of them have traditionally been answered through conducting sample surveys of the general public. The usefulness of social media depends therefore not just on whether it can answer these questions, but whether they can improve on existing methods in some respect. Where appropriate, we will also highlight key differences between social media data and traditional survey methods. In general we argue that, when compared to traditional surveys, social media data offer considerable advantages in terms of how quickly results are delivered, the scale at which results can be brought in, and (potentially) how cheaply they can be obtained. They also offer the possibility to access sub-groups within the population in a way that sample surveying has struggled with. The major difficulty lies in making accurate generalisations from social media data to some overall population of interest as those using social media (and especially those who use it frequently) do not constitute a representative sample of the public as a whole and do not come with perfect demographic data attached. This is an issue where much further research is needed.

2.2.1 Predicting the Present: Using social media to understand current salient issues

Knowing what the public is thinking about is a crucial precursor to knowing what their opinion is of any given topic. It is also an area where social media has the potential to offer real added value. A crucial problem in current opinion poll research is that while there may be a political need to know the public's opinion on a specific matter, the public themselves may have given the subject little attention. The European Union is a good example here: while it frequently attracts attention in the media and statements from politicians, polling often shows it to be a subject that the public think is of little importance. Hence asking for public "opinion" on the subject is almost meaningless, because a majority of people will not have been thinking about it before being asked the question. Some polling firms have tried to correct this problem by asking deliberately open ended questions, such as "what issues are the most important facing the country today?"[18] These questions however have the disadvantage of often giving very vague and generalist answers (for example, "the economy"). It is also difficult to be certain that people were

[18] See, for example, Ipsos MORI's "issues index": http://www.ipsos-mori.com/researchpublications/researcharchive/poll.aspx?oItemId=56

genuinely thinking about these issues, or whether they just think they are important when asked.

Offering an insight into currently salient issues is hence an area where social media has the potential to really fill in a gap. By providing a forum for unsolicited public comments and conversation to emerge, different social media platforms provide an indication of what the wide body of social media users are thinking about at any given time. It is no surprise therefore that a variety of indicators from social media are already starting to enter common parlance. For example, newspapers routinely report that a topic is "trending" on Twitter.[19] Other types of social generated data, such as traffic statistics in Wikipedia or search terms entered into Google, also offer the potential for knowing what people are thinking about even if they do not necessarily want to express or communicate with their friends. As we argued above, the usage rates of Google in particular are so high that they can genuinely claim to offer a type of overview of what a majority of people are thinking about (or at least looking for) at any given time, on the basis that if you are thinking about something you are quite likely to search for information on it. Hence a variety of researchers have used it as a type of broad measure of public attention.[20]

Public attention to different issues is important in and of itself. However researchers have also started to explore the potential this type of data has for predicting human behaviour, on the basis that informational searches often precede a particular activity. These types of methods are highly short term, and have hence sometimes been described as ways of "predicting the present".[21] Nevertheless they are of use because of the speed with which they can be delivered. Hence researchers have shown Google, Wikipedia and other types of social media to offer highly accurate predictions of (amongst other things) box office takings,[22] flu outbreaks,[23] stock market movements[24] and a whole variety of other socially interesting questions.

Finally, exploring the dynamics of public attention can also be a way of identifying key information sources which both inform people of what is going on and (potentially) help shape their opinions. This is something which traditional public opinion polling has always struggled to give a clear answer to. Of course, traditional surveys can ask questions such as "where do you

[19] The way Twitter's trends are determined is a commercial secret which is hence not fully understood. However broadly speaking Twitter looks for particular topics which have been part of a high volume of tweets over a short space of time.
[20] Ripberger, J. (2011). "Capturing Curiosity: Using Internet Search Trends to Measure Public Attentiveness.". *Policy Studies Journal*, 39(, 2),, 239-259.
[21] Choi, H., and Varian, H. (2012). "Predicting the Present with Google Trends.". *Economic Record*, 88, 2-9.
[22] See Mestyán, M., Yasseri, TahaT., and Kertész, J. (2013). "Early Prediction of Movie Box Office Success Based on Wikipedia Activity Big Data.". *PLoS One*, 8(8)
[23] See "Google Flu Trends", http://www.google.org/flutrends/
[24] See Preis, T., Moat, H., and Stanley, H. (2013). "Quantifying Trading Behavior in Financial Markets Using Google Trends.". *Scientific Reports*, 3., 1684

look for information on X". But the answers again can be quite vague: for example "newspapers" or "discussion with friends". Social media data, by contrast, offer the potential to pinpoint *which* newspaper (and indeed which article in that newspaper) or *which* friend.

The data on information sources is again made available by Google, who will report, for example, the top websites which are returned when someone inputs a particular search query. When these data are combined with information on the most popular search queries for a given topic, a powerful picture can be built up of the information landscape on offer to the public on a particular issue. More conversational social media such as Facebook and Twitter provide another angle. By looking at the type of links people post when they are talking about particular topics, we can see both who is talking about something and what information sources they rely on. The people talking about a certain topic can also be a valuable resource in and of themselves: a variety of studies have shown how social influence operates in subtle but effective ways in social networks to distribute information and inform people of what is going on.

There are two major reasons why any given organisation might want to understand the information landscape connected to a particular issue. Firstly, they may be interested to know where their own website is ranked when it comes to a certain mix of search terms, and thus their chances of attracting visits from people interested in particular topics (their own server logs should also provide a wealth of information on this). Secondly, they may also be interested to know which other organisations are influencing the debate on particular topics, and what their opinions are on the subject. This may help decisions be taken over, for example, how much effort to invest in relations with traditional news media, as quantifying the visibility of these media is now possible.

2.2.2 Social Media as a Source of Public Opinion

A second area of application of social media data to research purposes lies in the area of capturing public opinion on specific topics: i.e. knowing not only what the public are thinking about, but what they think about it. In this area, of course, existing social research tools such as the sample survey (but also other techniques such as focus groups) are more developed, while some of the problems inherent in social media data become more limiting. In particular, social media data allows us access to the opinions of the group of social media users who have chosen to express themselves in "public" (to the extent that social media are public, which of course varies by platform) on a particular topic. This group of people is very unlikely to be a random sample of the population, consisting not only of social media users (young, relatively technology literate) but also people with firm opinions on the subject who enjoy expressing themselves publicly.

Such sampling problems can be overcome, though methods for doing so are still developing. Many people report a wide variety of social characteristics on social networks (for example, age, gender, and location[25]): and these characteristics could be used to construct more representative samples than might be achieved by simply selecting users at random. Furthermore, techniques are developing for the automatic identification of such characteristics in cases where they are not actually available.[26] For example, identification of someone's gender has been shown to be possible based on their style of language use,[27] whilst someone's political preferences may be revealed through who they choose to connect to on Twitter.[28] Of course, social media will never be useful for accessing the opinions of people who do not use the internet; but previous surveying methods have also suffered from similar problems (for example, telephone surveys struggle to contact people who are ex-directory, whilst face-to-face interviews are biased towards those who are at home during the day).

In addition to correcting for potential bias in sampling, a further methodological problem here lies in the automatic detection of opinions or emotions expressed in messages. These questions can be addressed through a family of techniques often described as "sentiment analysis", which work in a variety of ways.[29] A simple example would involve constructing a list of words with a sentiment "score" attached to each one of them, and then assessing the extent to which each word is present in a given message. For example the existence of the word "hate" in the text might imply a negative message. A variety of more complex methods have been implemented in practice, which involve looking not just at the words but also groups of words, as well as the grammatical structure of text. However these techniques are still developing and significant challenges remain (for example, detecting jokes, irony or sarcasm is a constant difficulty). While a variety of academic and commercial packages exist for this task, none of them yet claim to offer a perfect solution.

Despite these limitations, the possibilities are also considerable. Public opinion polls take at least a few days to carry out and then synthesise into a report. Social media offers the promise of measuring "instant" reactions, which may give us a much more nuanced picture of how the public responds to breaking news. For example, a recently published study suggests that it is

[25] See Graham, M., Hale, S. A., and Gaffney, D. (2013). "Where in the World are You? Geolocation and Language Identification in Twitter.". *Professional Geographer*, forthcoming.
[26] See Kosinski, M., Stillwell, D., Graepel, T. (2013). "Private traits and attributes are predictable from digital records of human behavior.". *PNAS*, 110(15), 5802-5805
[27] See e.g. Newman, M., Groom, C., Handelman, L., and Pennebaker, J. (2008). "Gender Differences in Language Use.". *Discourse Processes*, 45(3), 211-236
[28] Golbeck, J., and Hansen, D. (2013). "A method for computing political preference among Twitter followers.". *Social Networks*, 36, 177-184.
[29] See Feldman, R. (2013). "Techniques and Applications for Sentiment Analysis.". *Communications of the ACM*, 56(4), 82-89

possible to measure the "mood" of Twitter users in relation to specific events. The researchers detected feelings of outrage and hate following the murder of Lee Rigby in London in May 2013; but also detected a mellowing of these feelings once the man's family appealed for calm.[30] Such immediate emotions which change very quickly would have been very difficult to detect using conventional polling tools which take more time to set up and deploy.

Furthermore, it is also important to note that sampling difficulties are only really relevant in cases where knowing the opinion of the *entire* population is the desired outcome. In many cases, however, simply knowing the opinion of social media users can be relevant, regardless of how representative they are. For example, businesses may be interested in the opinion of Twitter users because they represent a significant portion of the country and hence a significant market: whether this opinion can be generalised to anyone else does not affect its intrinsic importance. Furthermore, public networks such as Twitter offer extremely good coverage of certain social groupings, such as celebrities or journalists (which perceive them as an important part of their work) or young people. Or, on a narrower scale, specific social media such as educational sites or themed sites such as Mumsnet offer the promise of accessing specific communities. People interested in the opinions of these specific groupings may find the use of social media for opinion mining to be comparatively easier than, for example, trying to randomly survey a subsection of journalists; and they may even measure engagement in such communities as a useful measure of individual behaviour. For example, it has been suggested that students who are less involved on school social networks are more at risk of dropping out later.[31] Here social media can also be used in combination with more traditional survey techniques. For example, some studies have used Facebook groups or Twitter followers as a means of reaching out to specific sub populations and then actively soliciting their opinions using a more traditional survey based device[32].

2.3 Using Social Media Data: Challenges

Building on the discussion above, this section aims to review the practical challenges involved with using social media data, in two sections. First we explore technical challenges involved in social media data collection, and explore further some general methodological questions which have been raised in the previous sections. Secondly, we look briefly at legal and ethical concerns inherent in using socially generated data.

[30] For more details see: http://emotive.lboro.ac.uk/
[31] Wright, F., White, D., Hirst, T., and Cann, A. (2014). "Visitors and Residents: mapping students attitudes to academic use of social networks.". *Learning Media and Technology*, 39, 1, 126-141.
[32] Bartlett, J., Birdwell, J., and Littler, M. (2011). *The New Face of Digital Populism*. London: DEMOS.

2.3.1 Technical and Methodological Challenges

Using social media data for any purpose presents first of all a series of important technical challenges. At a general level, significant computer skills are required both in terms of gaining access to the data and using it for analysis, skills which are currently in quite short supply. For organisations this implies either hiring skilled staff, or contracting out to a private consultancy which specialise in social data analytics, Off-the-shelf solutions do exist, but offer limited insight for those without sufficient awareness of how the internals of the system work.

Furthermore, organisations wishing to engage in their own analytics need a significant IT infrastructure to support them. Engaging in projects such as measuring the mood of the nation on Twitter involves a computer capacity which is capable of both collecting and storing significant quantities of data. The costs of storage are constantly decreasing, meaning that this infrastructure is not out of reach of small and medium sized organisations. But the costs involved are nevertheless worth considering.

A further challenge is also posed by the fact that social media data are easiest to collect at the moment of their creation. The legal controls surrounding access to such data are discussed more fully below, however typically services such as Twitter do not make entire archives of tweets available: rather, information can be accessed only as it is being published (though it is also possible to retrospectively purchase data through commercial suppliers). This means a degree of rapid reaction is necessary with any research project. For example, when researchers at the Oxford Internet Institute were studying the development of the Arab Spring on Twitter they needed to react very quickly to events which could not be foreseen in advance, to make sure information was captured as it was created and hence avoid expensive commercial solutions.[33]

A final and significant problem for social media research concerns the extent to which sentiment measured on such networks can be attributed to “real people”. As a result of the commercial value and significance of social media, a number of professional organizations exist which try to actively influence overall perceived sentiment. This occurs both through active engagement with the communities on these networks, and (more problematically for our purposes) through the creation of fake accounts which are designed to rebroadcast messages, to swell the friend / follower count of certain individuals, or otherwise make the overall social media landscape look different than it would do “naturally”. This process, sometimes defined as “astroturfing”, has been reported by students at the OII in the US presidential election (when Mitt Romney’s Twitter follower count was suddenly and

---

[33] See Gonzalez-Bailon, S., and Wang, N. (2013). "The Bridges and Brokers of Global Campaigns in the Context of Social Media.". Available from: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2268165

dramatically inflated).[34] Again, this is a problem that can be counteracted: methods exist for the discovery of fake accounts, and social networks themselves are quite active in trying to find and delete them (Facebook, for example, recently deleted millions of apparently fake accounts).

Finally, it is worth offering a brief discussion of many of the methodological challenges highlighted in section 1.2: the fact that social media users are not representative of the general public, that only a minority of users post the majority of content, that opinions which are posted are unsolicited, and that accurately assessing these opinions in an automatic way is difficult. We have discussed above a variety of methods for counteracting each individual problem; here it is worth mentioning the major general way of counteracting all of them, which is “benchmarking” social media measures. Benchmarking can be conducted in a variety of ways. We can compare two different concepts of interest at the same time, or we could track the same concept over a period of time, looking at whether a particular policy is getting more or less popular. Finally, we can link social media data to other data sources which are more known and trusted; for example, we could link the social media “score” for a given agency or service to other metrics of outcome and performance. At this relatively early stage in the development of social media research, these different types of benchmark measures are crucial for adding meaning to otherwise quite abstract figures, and instilling some trust in the overall method.

2.3.2 Legal and Ethical Challenges

There are also significant legal and ethical concerns around the access of social media data, concerning the relationship of researchers both to the company which created the platform for the data and which hosts it on its servers, and to the end users which actually contributed to the content. These concerns are reviewed here in a general way.

Firstly, the use of social media data creates a number of obligations towards the company which owns that data, which is typically the one which also developed the platform on which the data are then created (hence Facebook owns all of the data created on Facebook). The exact nature of these obligations is still evolving. In recognition of their increasing use in research (and more generally by third party applications), many of the larger organisations have summed up these obligations in specific “Terms of Service” [TOS] which are usually accessible on each organisation’s website. While each TOS differs in terms of content, many organisations are typically happy for data which is publicly available to be collected providing it is not

[34] See Furnas, A., and Gaffney, D. (2012.). “Statistical Probability That Mitt Romney's New Twitter Followers Are Just Normal Users: 0%”. *The Atlantic*. Available from: http://www.theatlantic.com/technology/archive/2012/07/statistical-probability-that-mitt-romneys-new-twitter-followers-are-just-normal-users-0/260539/

made available for download elsewhere. Many of these organisations also provide specific technical interfaces for the accessing of data, known as “Application Programming Interfaces” [API], which allow them to both monitor and set limits on it (a typical Twitter researcher will, for example, only be able to access 1% of the material published on Twitter on any given day).

Smaller social sites may also provide terms of service, but are less likely to provide API access; hence data must be downloaded by automatically instructing a computer to extract relevant information from the web pages, a process sometimes known as “scraping”. Such scraping lies in a legal grey area as, while the content is provided for free, it is also typically protected by various copyright and intellectual property laws. Furthermore, scraping also involves repeatedly accessing a given website, which if done on a large scale may place an unreasonable burden on the web servers of these sites. The legal landscape in this area is again still evolving, hence it is difficult to say definitively what is legal and what is not: however researchers which obey TOS documents, minimize server load, use data for non-commercial purposes and do not make it available for re-use elsewhere are again likely to avoid problems.

Finally, significant amounts of social media data lie beyond the reach of even scraping techniques. Facebook is a good example: while it is possible to automatically access data on your own Facebook account, collecting data *en masse* from Facebook’s servers is possible only for a limited set of data types. Facebook itself has started to manage its interactions with researchers through its own in house team of sociologists: hence people working with Facebook data may also be required to work with Facebook’s own researchers.

Beyond obligations to the data owners, social media researchers also have a series of obligations towards the people that created the data originally. Again, law in the area is developing, and is also complicated by the fact that these people may come from all over the world, hence are theoretically subject to overlapping legal regimes (in Europe, the most relevant piece of legislation is the European Data Protection Directive). In general they should think relatively carefully about how they store data, especially in terms of how secure that storage is, and do their best to present results and data in an anonymous / aggregate fashion. Of course, a variety of studies have demonstrated that it is frequently possible to “de-anonymize” data which had previously been thought to have stripped of all personally identifying information; as many seemingly less personal bits of information can nevertheless be linked to individuals if enough of it is available.[35].

In ethical terms, researchers face the problem that the classical approach to ensure ethical participation in a given study, which is based

[35] See Sweeney, L. (2002). "*k*-anonymity: a model for protecting privacy.". *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, 10(5), 557-570.

around the idea of informed consent, is essentially impossible, both because of the sheer numbers of people involved and the practical challenge of contacting them. While these people are theoretically aware when they create content that it may be both made public and used for other purposes, in practice few will have considered that it might one day become the object of a research project.

For this reason, most contemporary researchers base their research ethics around the idea of “minimizing harm”, which revolves essentially around making sure that people whose data are utilized do not suffer any negative effects from this utilization. In practice this boils down to techniques quite similar to those when minimizing legal concerns around data protection: making sure the data are anonymized and stored securely, and that only aggregate level results are reported.

3 Part 2 –Social Media Data and DWP Policy Research: the cases of Universal Credit and Personal Independence Payment

In this second section, we move on to an empirical exploration of the potential of social media for DWP work, looking in particular at how it might be used to support policy relevant social research on two current policy projects: Universal Credit [UC] and Personal Independence Payment [PIP]. We have chosen within the context of this relatively short feasibility study to focus on the broad scope “mass” social media described above (such as Twitter and Facebook), as these media offer comparatively greater potential for data extraction over short time periods, and as they form the main focus of some of the current excitement around social media analytics. However we also think there is significant potential in some of the more narrow scope social media forums we mentioned (such as Mumsnet), which could be explored in subsequent projects, a point we return to in the conclusion.

The empirical work presented here is divided into four sub-sections, each one focusing on a different type of social media data. In each section, we look at how the two key uses of social media data identified in section 1 (ways of measuring public attention to different issues and ways of measuring opinion on those issues) may provide insight for social researches at the DWP.

In section 3.1, we explore data coming from Google. As discussed above, Google provides a significant amount of information on the behaviour of its customers (“socially generated data”), much of which can be harnessed to tell what people are thinking about, and when they are thinking about it. In this section we explore some of these possibilities.

In section 3.2 we move the focus to Wikipedia, probably the most important general source of online information. We look at how the amount of visits to different Wikipedia pages has evolved over time, and how the structure of the pages themselves has been changed, as further indications of public awareness and attention to different welfare policies. We also construct certain measures of page quality, and argue that DWP should pay attention to Wikipedia.

In section 3.3 we look at the broad concept of “mentions” on a range of social media platforms: individual postings to social media sites where a particular word or phrase of interest appears. We explore the extent to which this may be an indicator of both public attention to a particular policy issue and broader public awareness of that issue.

In section 3.4, finally, we focus more specifically on the DWP’s own social media presence, exploring the activity surrounding almost 500 of the DWP’s Twitter accounts. We look in particular at the use of automated sentiment analysis techniques to explore feedback coming into these different branches, and examine whether this is a useful way of identifying particular problems or challenges that they may be facing.

A brief note on data collection methodology. Social media data is often available directly from social media platforms via Application Programming Interfaces (APIs), which are interfaces created by these platforms designed to facilitate data capture and re-use. Whilst websites are designed for human use and mix presentational elements with data, APIs generally focus purely on data and are designed to be easily read by machines. A computer script can request data from an API following the documented standard, and the API will respond with the requested data in a machine-readable format. Most social media platforms limit the type and volume of data that can be obtained from their APIs. To overcome these limitations commercial companies buy, aggregate, enhance, filter, and sell raw data (e.g., Gnip, Datasift, Topsy) and/or offer basic, automated analysis tools (Hootsuite, Radian 6, Brandwatch, etc.).

The data for this section of the study comes from a mix of direct-to-source API data and commercial data companies. The data for YouTube, Facebook, and Google was collected directly from each provider via the i2-sm data collection platform.[36] The platform was also used to gather Twitter data, which was sourced from one of Twitter’s commercial data partners, Topsy, and augmented further with information directly from the Twitter API. The i2-sm platform allowed for consistent scheduling of data collection queries as well as augmentation and storage of the data.

A list of keywords was developed to identify content possibly relevant to the study (the full list is available in the annex). These keywords included terms like “dwp,” “Universal Credit”, “JSA”, “Jobseeker’s Allowance” and “PIP”. The initial keyword list was designed to be very broad as it is much easier to discard irrelevant content later than to obtain historical data from most platforms. There are several techniques to “disambiguate” or separate

[36] i2-sm (http://www.i2-sm.com/) is a London technology start up with which one of the report’s authors is involved.

different uses of terms like “JSA,” which could be used for Jobseeker’s Allowance, the Japan Sumo Association, or Job Services Australia based on the surrounding content. Given the pilot nature of this study, we discarded results that only included an ambiguous term like “JSA,” and our analysis proceeds with longer, unambiguous strings only. A fuller, long term study should try different disambiguation techniques and evaluate their performances with human analysis.

Data was collected from Twitter, YouTube, and Facebook from 20 March 2013 to 30 September 2013. YouTube and Twitter were checked once per week and any content in the last week matching one of the study’s keywords was downloaded. Facebook was checked daily through its public search API for any public posts matching one of the study’s keywords. Google search results were recorded once per day from 8 December 2012 to 30 September 2013 in the same way. Further data was obtained directly by the authors at the time of the study from Wikipedia for age view statistics, from Google for search volumes and related queries, and from Twitter for information on the DWP’s own accounts.

3.1 Search Data

Search engines, such as Yahoo, Bing and (especially) Google, are a vital part of the way most people use the internet. The overwhelming majority of internet users use search engines at least some of the time, while a majority (around 60%) use them as their only major way of finding things online.[37] Search engines are amongst the most visited websites in the world, and record staggering amounts of search queries (in 2012, for example, Google recorded around 115 billion search queries, approximately 65% of the global market share[38]). Beyond its sheer scale, the importance of search as a data source lies not so much in its ability to capture people’s opinions or preferences (though, especially with the rise of the GooglePlus social network, this is something that Google is increasingly building a picture of as well), but rather its ability to tell what people are thinking about and when they are thinking about it.

All search engine use is about finding information. However work on search engines typically distinguishes between three types of use, depending on the specificity of the information sought:

- Informational queries where people are searching for general information, for example about the types of benefits they might be eligible for

[37] Dutton, W., and Blank, G. (2011), 22

[38] Sullivan, D. (2013). "Google Still World’s Most Popular Search Engine By Far, But Share Of Unique Searchers Dips Slightly". *Search Engine Land*. Available from: http://searchengineland.com/google-worlds-most-popular-search-engine-148089

- Transactional queries where people are looking for a place to conduct a specific action, for example where to apply for a specific benefit
- Navigational queries where people already know the website they want to visit but use Google to find its specific address, for example searching for “Universal Jobmatch”

There are, we suggest, three major uses to which the DWP could put this type of data in the context of social research on UC and PIP. Firstly, Google presents the opportunity to explore the extent to which people are aware of the policies in question. Both policies represent changes to the system whereby benefits are claimed in the UK. The extent to which people are aware of these changes could prove important for the extent to which they are adopted successfully. This is explored in section 3.1.1. Secondly, Google data may also be usable in improving short term awareness of when the population are thinking of specific topics of interest which may indicate important forthcoming changes in behaviour. The DWP may want to know, for example, if, in a specific area of the country, searches for the word “ available jobs ” have experienced a significant growth - this may indicate a coming change in the amount of benefits claimants. This is something explored in section 3.1.2. Finally, Google data provides a picture of how a large segment of the population finds information about topics with which they are unfamiliar. This may be useful in terms of conducting research on the "sources" of public opinion: and will also allow the DWP to examine its own position in the information landscape surrounding a policy. This is explored in section 3.1.3.

3.1.1 Google Trends and Public Awareness of DWP Policies

In this section, we explore the use of Google data as a proxy for public awareness of our key policies of interest (Universal Credit and Personal Independence Payment). This information, we argue, is an important research topic in its own right, especially in the context of newly implemented policies: as better public awareness may improve implementation and adoption of new proposals.

We explore these questions using the Google Trends service.[39] This service provides a means of comparing the frequency of use of a search term over time, and also allows us to compare multiple terms together (it does not, however, provide absolute numbers of searches). This comparison is useful for benchmarking purposes, as a way of exploring the progress of a new policy initiative. Clearly, as Universal Credit is itself rolled out, one of the aspirations must be that it “replaces” earlier terminology such as jobseeker’s

[39] See: www.google.com/trends

allowance in the minds of the general public. The more the public have a clear idea what the service they are interested in is called, the easier it will be for them to communicate with the DWP and to find information on it in general.

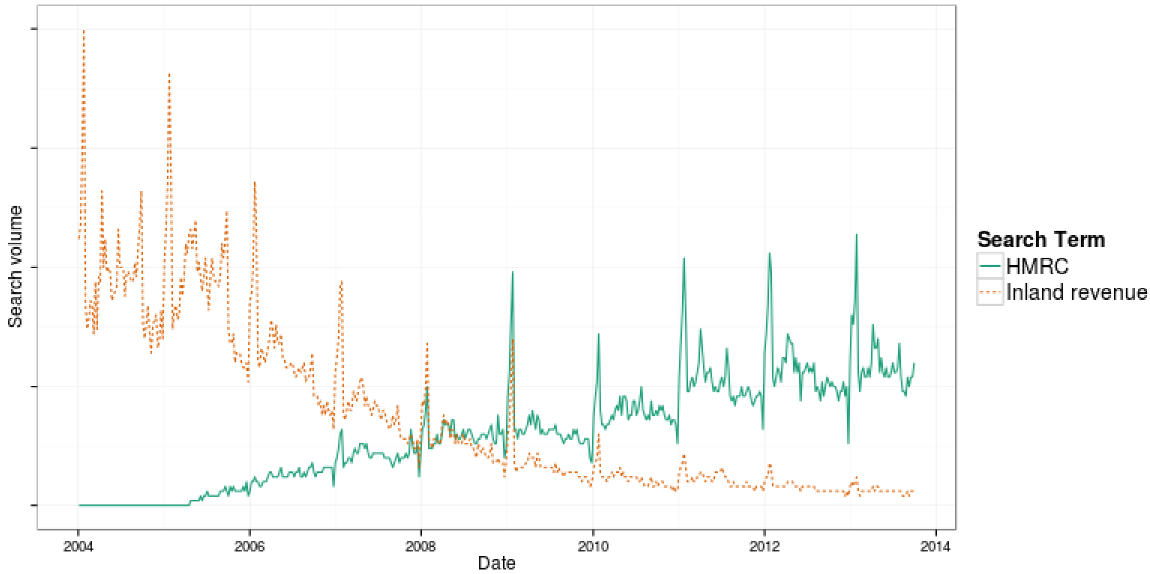


Figure 1: HMRC and Inland Revenue compared as Google Search Terms

Google trends provides a useful means of testing the extent to which any given term is “replacing” another. By way of example, we can use the replacement of the Inland Revenue with the body Her Majesty’s Revenue and Customs [HMRC] in 2005. Figure one compares the search terms “HMRC” with “Inland Revenue”. It shows clearly two things. Firstly, people search for both of these terms more at certain points of the year. Secondly, the frequency of use of the two terms crossed over only in 2008; nearly three years after HMRC had been formed. Only from 2011 onwards can we really see searches for the old term dying out.

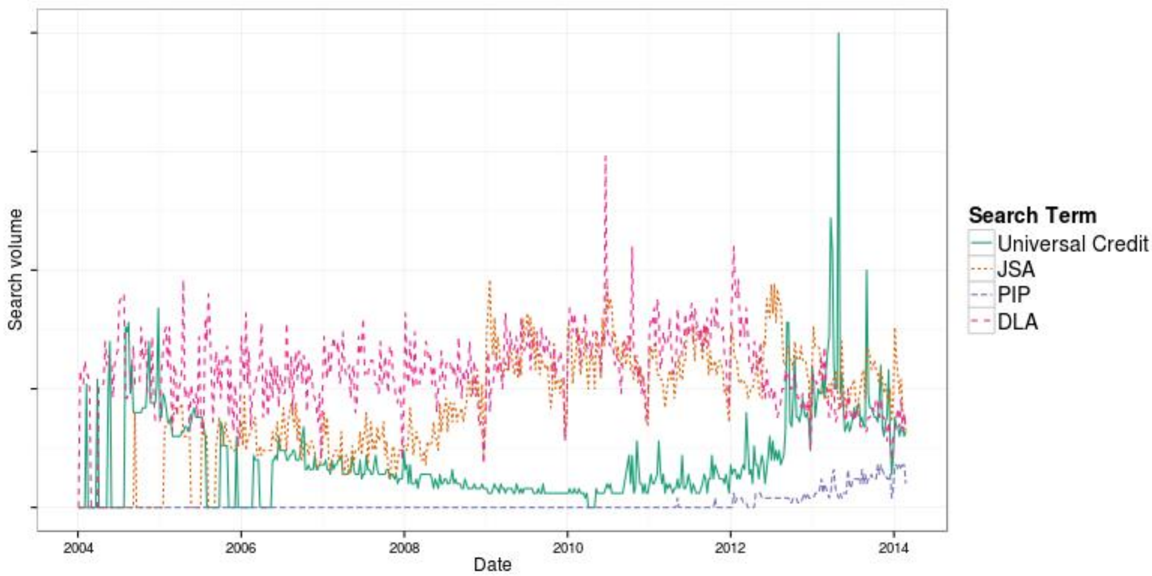


Figure 2: Universal Credit, Personal Independence Payment and Jobseeker's Allowance as Google search terms. Highest peak represents approximately 60,000 monthly searches[40]

It is useful to compare this example with the experience of Universal Credit and Jobseeker's Allowance, which are compared in figure two (we also include PIP and DLA as search terms for the purposes of comparison). We can see that searches for the term Universal Credit began to increase almost as soon as it was announced in October 2010, but it was not until 2013 that it reached comparable levels of awareness with jobseeker's allowance in terms of amount of searches made. It has certainly not replaced jobseeker's allowance as a search term yet (which continues to be at a relatively high level), though we would perhaps not expect this given that it has not been fully rolled out.

It is also worth noting here that, despite a much larger claimant base the volume of searches for PIP is considerably lower overall than both JSA and Universal Credit. There are several potential reasons for this. One is that press coverage of Universal Credit has been considerably higher, which may drive public interest in the topic. Another is that "Personal Independence Payment" as a search term may not be the most common way of searching for this type of benefit. It may be that different underlying dynamics in the reasons for claiming these benefits produce differing behaviours on Google Search (for example, people may become redundant unexpectedly, prompting them to search for information). Finally, it may be reflective of the underlying profile of internet users, which may under represent PIP claimants or over represent those claiming JSA.

[40] Google typically disregards punctuation marks such as apostrophes, hence the search term entered is "jobseekers allowance" rather than "jobseeker's allowance"

The potential permutations of keywords which people might use to search for things of interest to the DWP are of course almost endless. However, Google Trends also provides a way of discovering how different people use different keyword mixes through its “related searches” feature. This is a useful way of discovering other search terms which may be of interest for DWP work. Figure three below, for example, shows that jobseeker’s allowance is related to a variety of other very similar searches, whereas Universal Credit is connected to ideas of both “tax” and “benefit”.

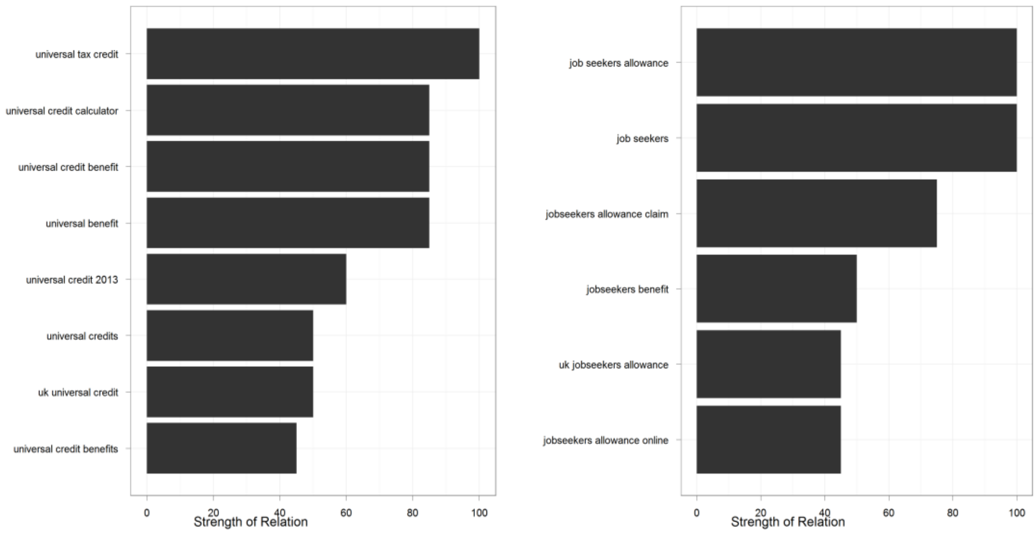


Figure 3: Top Related Searches for Universal Credit (left) and Jobseeker’s Allowance (right)

Finally, it is worth noting in this context that Google does offer a paid service, AdWords, which provides direct access to customers searching for particular keyword mixes (such services are also employed by companies which specialise in providing advice to benefits claimants[41]). Though our overall recommendation would be to focus on free ways of improving search ranking, it may be worth considering experiments with small amounts of paid advertising to see if this is a more effective way of funnelling search engine traffic. AdWords also provides a lot of information about the relative popularity of different keywords, and the extent to which there is competition in the market on such words.

3.1.2 Using Google Trends to "predict the present"

In this section, we explore the potential use of trends data as a means of “predicting the present”: providing a rapid awareness of developing situations which could be of interest to social researchers at the DWP. This can be achieved by exploring the extent to which a change in the raw numbers of people searching for a particular term could be linked to a real

[41] See in this context a recent report by the Advertising Standards Authority on “Copycat Websites” (available from: http://www.asa.org.uk/News-resources/Media-Centre/2014/Copycat-websites.aspx).

world outcome which the DWP is interested in. One example of this is provided by Google Flu Trends, a research project run by Google itself, which uses an increase in search terms for common flu symptoms (e.g. headache, sore throat) as an indicator of overall flu patterns. The results have been shown to provide a highly accurate picture of flu patterns in a variety of countries[42] (though their accuracy has recently become more open to question). Figure four shows some data from the flu trends project in the US, which records Google's estimate against real data taken from the US Centre for Disease Control. The two lines follow each other closely (and it is worth highlighting again that the Google data is available in real time, while the CDC data is available only after 1-2 weeks).

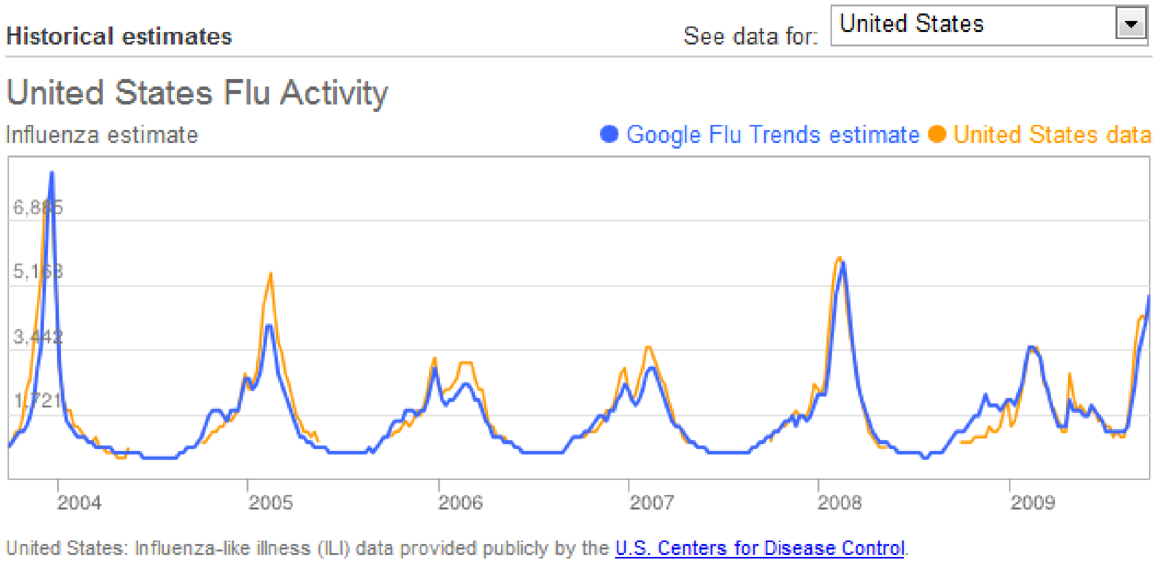
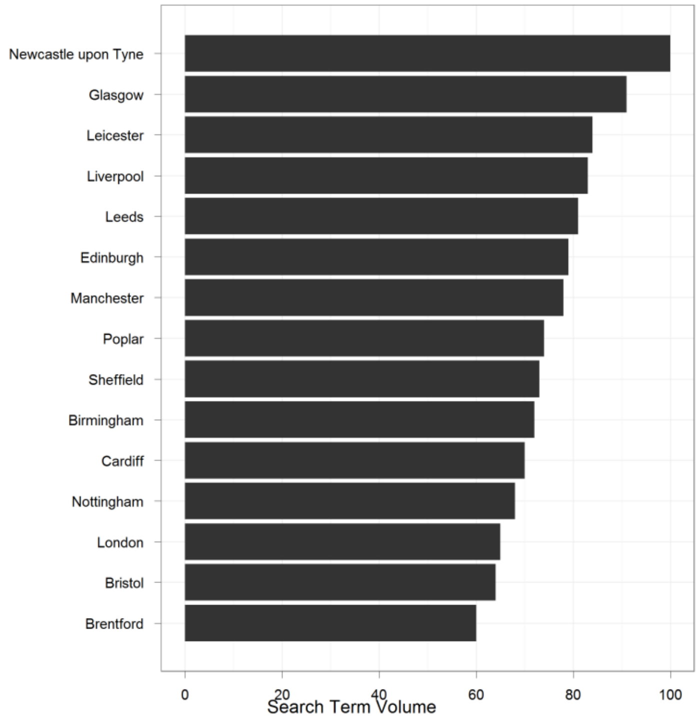


Figure 4: Google Flu Trends

There are a variety of ways which the DWP could consider using such data. A rapid rise in searches terms of interest in a particular area may translate several months down the road into increased activity at a local Jobcentre. As figure five below shows, at the time the research was conducted the most active city in terms of searches for jobseeker's allowance was Newcastle. The right hand side of figure five provides, for the basis of comparison, search term volume figures for the keyword "Facebook", which we would expect to have a relatively neutral distribution across cities.

[42] Ginsberg, J., Mohebbi, M., Patel, R., Brammer, L., Smolinski, M., and Brilliant, L. (2009). "Detecting influenza epidemics using search engine query data.". *Nature*, 457, 1012-1014



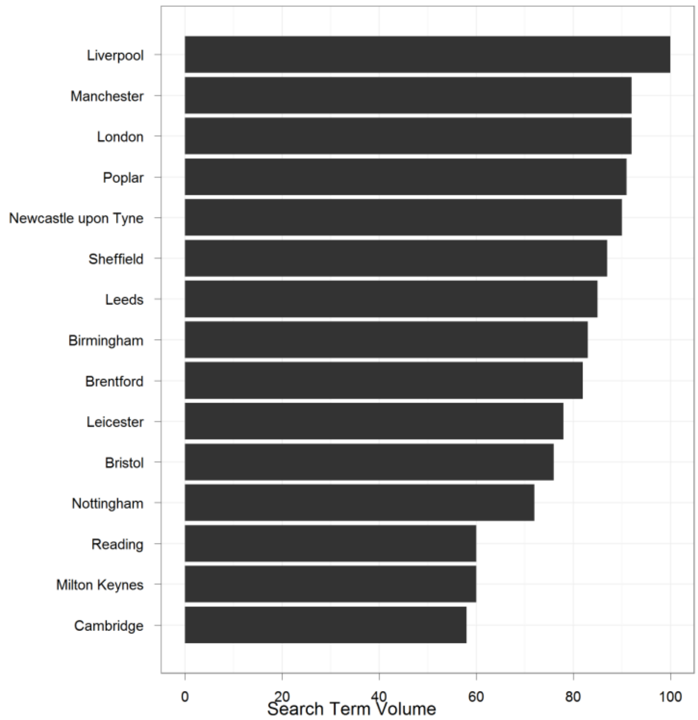


Figure 5: Volume of searches for Jobseeker's Allowance (top) and Facebook (bottom) in different cities

Within the scope of this short feasibility study, we cannot of course prove whether rises in this particular search term have any predictive power. A vital part of using such data for prediction benchmarking these terms against existing known data over time. The success of the Google Flu trends project is based on observing search terms over a long period of time and seeing the extent to which they are predictive of actual real world outcomes. Further research and exploration would therefore be required to find what keyword mixes, if any, are genuinely useful predictors of a coming rise in unemployment. However, as regional joblessness statistics are widely available, this is something which could certainly be tackled.

3.1.3 Google Search and the sources of public opinion

In this section, we explore the use of data taken from the Google Search API to explore key sources of information on the two policies of interest. As we argue above, knowledge of these information sources is useful because they may shed light on the roots of public opinion about a given policy. This knowledge also allows the DWP to shape its own information policy, in particular by identifying the extent to which its own websites are well represented in the information landscape.

The search API allows the user to enter a specific keyword or set of keywords, and then find out which websites are the top search results are for those keywords.[43] Figure six provides results for the search term “Universal Credit”. The graph displays the average rank for the seven most important websites over the period December 2012 to August 2013. So, for example, the website www.universalcredit.co.uk was on average the highest result reported at the beginning of April in that year.

This graph highlights a number of interesting findings. Firstly the information landscape is quite congested, with newspapers, NGOs, campaigning organisations, informational sources such as Wikipedia and the DWP itself all jostling for position. The rankings highlight quite clearly however *which* NGOs, newspapers, campaigning organisations and information sources are important (and a closer analysis of the results will also point to specific authors and articles within newspapers). This provides a clear list of organisations the DWP ought to pay close attention to as key opinion shapers. It is worth highlighting here that the top ranked site for much of the period (www.universalcredit.co.uk) is a private site which the DWP does not exercise any control over.

[43] We should note here that the results from the API are not identical to those which come out of Google Search, as search also uses a range of personal filters to try and customize its search results to individual people (based on, for example, search previous search behaviour, location, or language settings). Nevertheless overall these data provide a good impression of what’s popular.

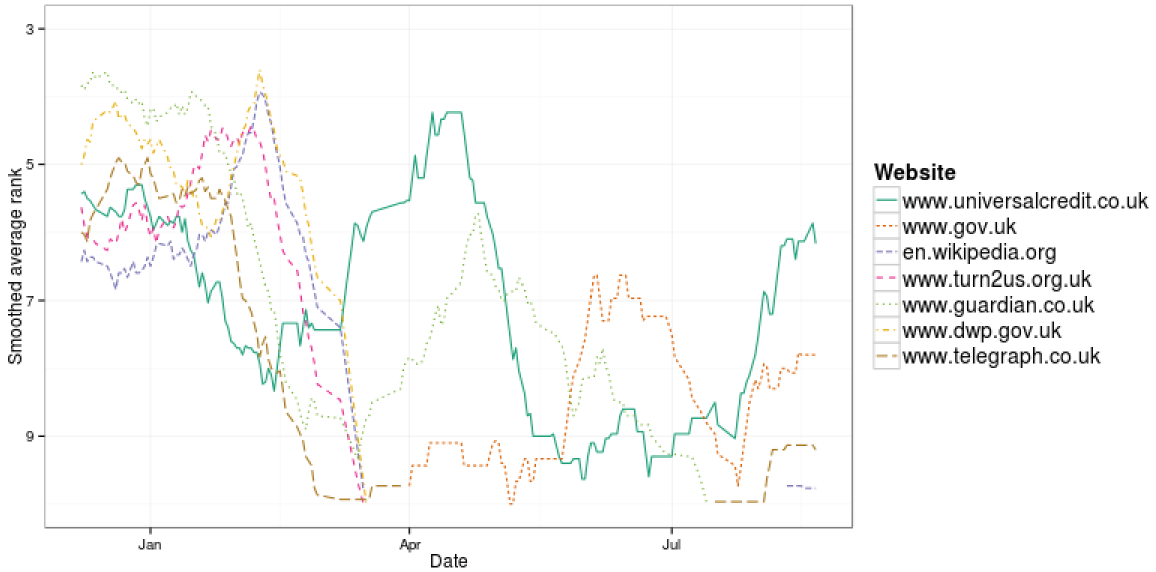


Figure *6*: Google API Results for the search term "*Universal Credit*". *December 2012 – August 2013.*

Secondly, if we focus just on official government websites (which are picked out in figure seven), we can see that for most of the sampling window they achieved a good position near the top of the ranking. However there was a considerable break in the period from March to June (coinciding with the move of information from the DWP site to the gov.uk site). We can also see that and the newer gov.uk website did not get back to the search ranking achieved by the previous DWP website during the window of observation (though it is worth nothing that at the time of writing its ranking appears to have improved on what is displayed below). As is to be expected this means that, on average, this website appears less frequently in a high search position than the previous DWP one, which makes it likely that less people visit it when searching for "universal credit" on Google.

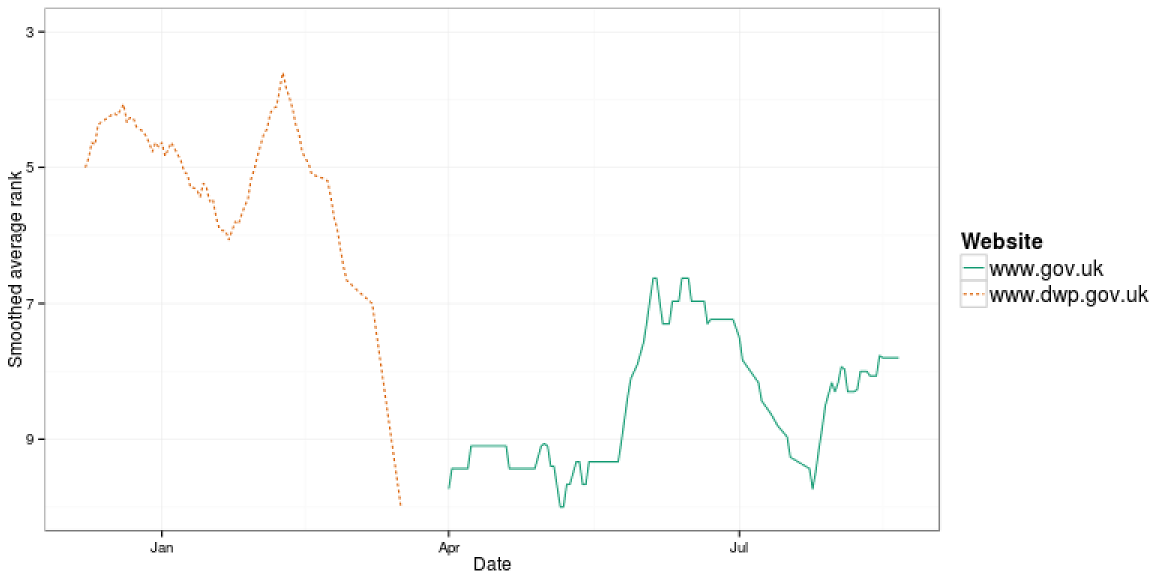


Figure *7*: Google API Results for the search term "*Universal Credit*", *limited to just gov.uk and dwp.gov.uk. December 2012 – August 2013.*

3.2 Wikipedia Presence

In this second empirical section we explore the representation of DWP polices on Wikipedia. Wikipedia is a web-based encyclopedia whose content is created and updated by its users. It is undoubtedly the most popular source of general information on the internet, with some studies ranking Wikipedia use as the third most popular online activity. Its popularity is driven in part by its strong Google search ranking: one study found that Wikipedia articles feature in the top 5 Google Searches around 96% of the time.[44]

Wikipedia is important in the context of policy research for several reasons. Firstly it acts as a vital source of information for the general public for a huge range of different topics.[45] The extent to which DWP relevant Wikipedia pages are viewed hence provides another proxy for the extent to which the public are paying attention to different projects of interest. Secondly, the importance of Wikipedia for informing the public means that the quality of the Wikipedia page on policies such as UC and PIP, and their overall position in the Wikipedia website, are factors of importance in terms of the shaping of public opinion.

3.2.1 Wikipedia Readership Statistics as an Indicator of Public Awareness

We begin by looking at patterns of readership of the Wikipedia pages for Universal Credit and Personal Independence Payment, as well as the page which refers to Jobseeker's Allowance for the purposes of comparison. Monthly readership patterns are set out in figure eight, with daily patterns in figure nine. These graphs highlight a number of findings. Up until approximately the end of 2012, readership figures for Universal Credit remained consistently between 1,000 – 1,500 per month. After July 2012, by contrast, they came up to a level similar to that of Jobseeker's Allowance, fluctuating around the 7,500 per month figure. Interestingly, these fluctuations then started to mirror that of Jobseeker's Allowance for the rest of the time period under observation. PIP, by contrast, receives a very small amount of traffic throughout the window of observation – a finding consistent with that produced in section 3.1.1.

[44] See Samoilenko, A., and Yasseri, T., (2014) The distorted mirror of Wikipedia: a quantitative analysis of Wikipedia coverage of academics, *EPJ Data Science*.

[45] See Margetts, H., and Hale, S. (2010) "User Experiments. Phase 1: Life Events Report". *Study on user expectations of a life events approach for designing e-government services*, European Commission, 79-96.

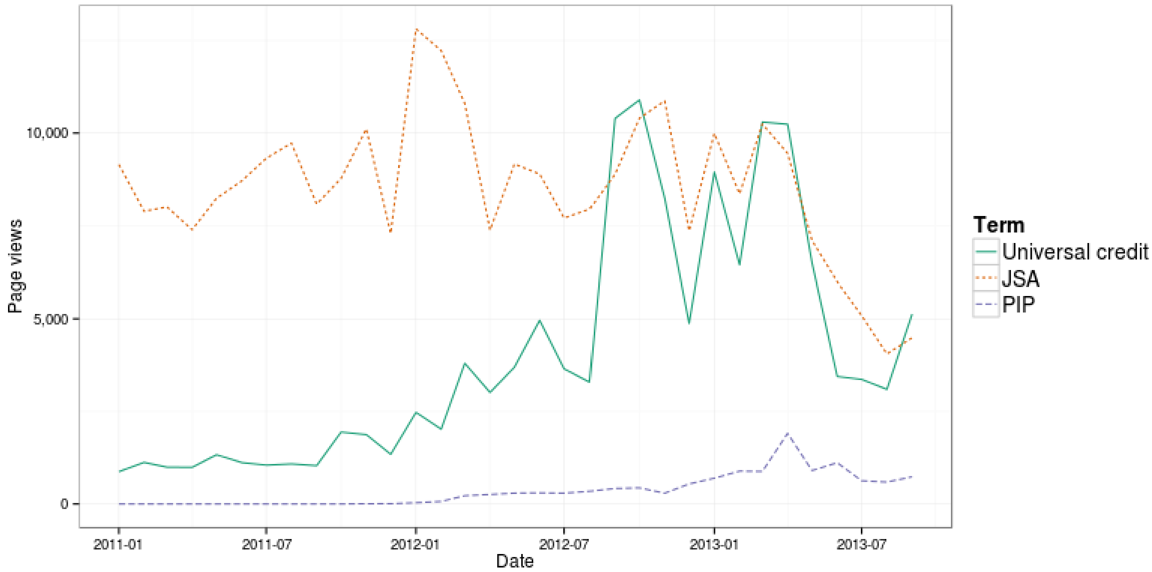


Figure 8: Volume of *monthly page views for three Wikipedia pages*

The volume of people visiting each Wikipedia page demonstrates the overall power of Wikipedia to influence people. It also offers an approximate measure of public attention and awareness which fits in with those observed in other platforms: Wikipedia views for Jobseeker's Allowance and Universal Credit start to mirror each other at approximately the time where the two search terms start to overlap on Google.

3.2.2 Wikipedia Article Quality as an Indicator of Public Awareness

A second angle on public awareness can be provided by the quality of the Wikipedia articles for UC and PIP. One way of getting a handle on the quality of a Wikipedia article is to explore its "edit history". These edits are created by Wikipedia's community of editors. As edits are often motivated by current news events the amount of edits gives an indirect impression of the extent to which a particular topic is in the public eye. The length of an article will also be an indication of the amount of work Wikipedia has put into it, with articles typically getting longer as they get older and more established.[46]

[46] Wilkinson, D., and Huberman, B. (2007). Cooperation and quality in Wikipedia. *WikiSym '07 Proceedings of the 2007 international symposium on Wikis.*

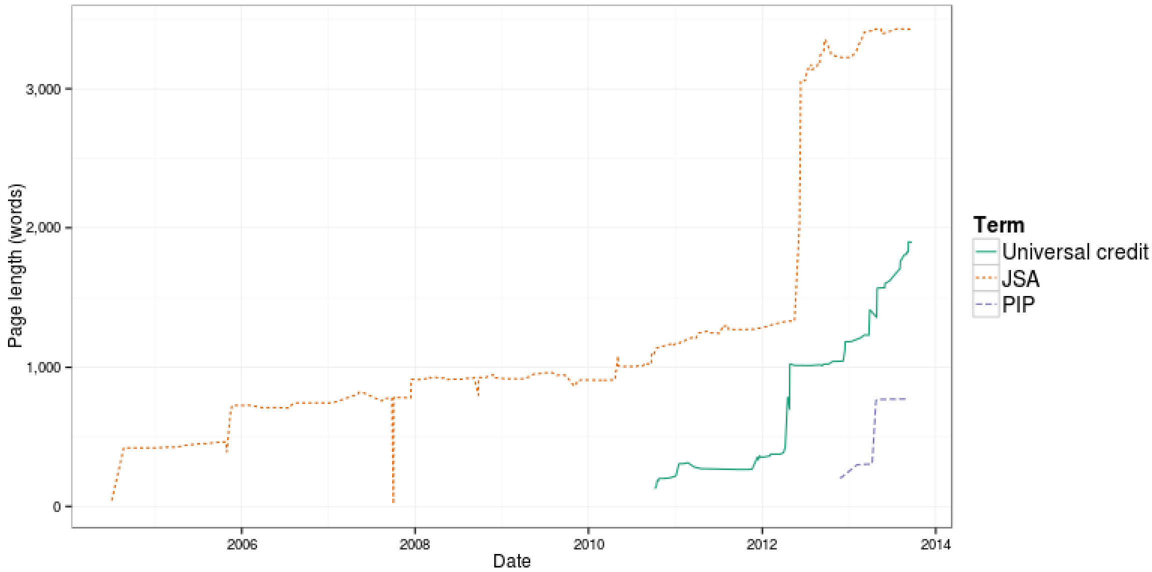
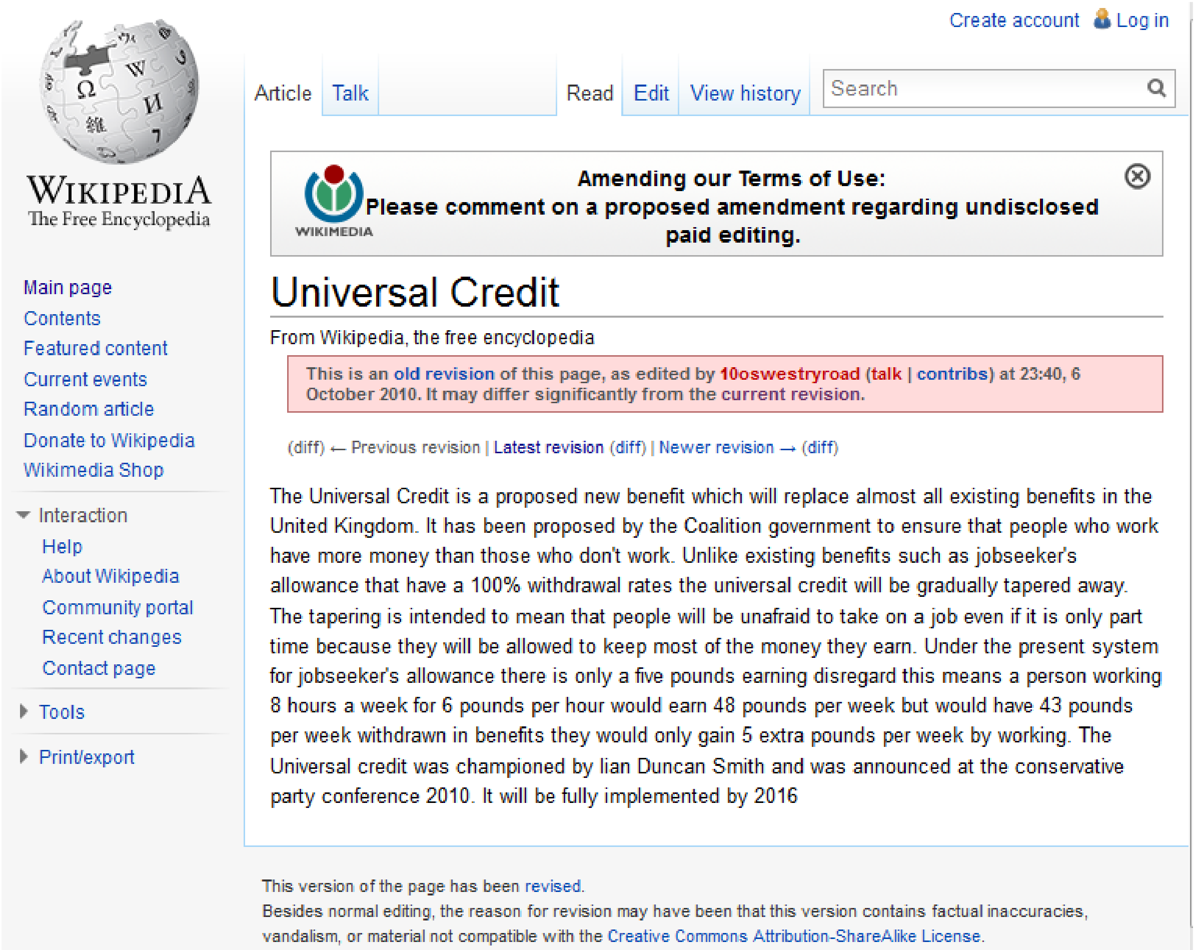


Figure 10: *Length of the UC, JSA and PIP Wikipedia Pages (in words)*

The amount of edits is tracked in both figure ten and table one. Figure ten explores in particular how page length has changed over time. We can see both that page length typically increases over time, but also that the majority of these increases happen in short sharp bursts. For Universal Credit and Jobseeker's Allowance, we can see that the key formative moment for the page was in February 2012 (much earlier than the spikes identified in other sections). For PIP this moment occurs about a year later. We can also see that, up until this moment in 2012, the UC page remained almost identical to the one which was first posted in October 2010, a copy of which is shown in figure eleven. This page describes the scheme very briefly in abstract, but contains little details on implementation, on the pilot schemes, or links to official information.



Create account Log in

Article Talk Read Edit View history Search

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikimedia Shop

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools

Print/export

WIKIMEDIA

Amending our Terms of Use:
Please comment on a proposed amendment regarding undisclosed paid editing.

Universal Credit

From Wikipedia, the free encyclopedia

This is an old revision of this page, as edited by 10oswestryroad (talk | contribs) at 23:40, 6 October 2010. It may differ significantly from the current revision.

(diff) ← Previous revision | Latest revision (diff) | Newer revision → (diff)

The Universal Credit is a proposed new benefit which will replace almost all existing benefits in the United Kingdom. It has been proposed by the Coalition government to ensure that people who work have more money than those who don't work. Unlike existing benefits such as jobseeker's allowance that have a 100% withdrawal rates the universal credit will be gradually tapered away. The tapering is intended to mean that people will be unafraid to take on a job even if it is only part time because they will be allowed to keep most of the money they earn. Under the present system for jobseeker's allowance there is only a five pounds earning disregard this means a person working 8 hours a week for 6 pounds per hour would earn 48 pounds per week but would have 43 pounds per week withdrawn in benefits they would only gain 5 extra pounds per week by working. The Universal credit was championed by Iian Duncan Smith and was announced at the conservative party conference 2010. It will be fully implemented by 2016

This version of the page has been revised.
Besides normal editing, the reason for revision may have been that this version contains factual inaccuracies, vandalism, or material not compatible with the Creative Commons Attribution-ShareAlike License.

Figure 11: The first revision of Wikipedia's article on Universal Credit, which appeared the day after it was announced by Iain Duncan Smith in October 2010.

Table one compares UC, JSA and PIP to other topics from contemporary UK politics, in order to provide some benchmarks (including as well average daily readership statistics). The articles are ordered by overall length, though this is closely related to the number of edits, the number of editors and the number of articles linking in. We can see that UC and PIP are at the lower end of the spectrum, both shorter and less well edited than many other government policies of the same era, even ones which have received comparatively less press attention such as the creation of Police and Crime Commissioners. The Universal Credit article is, for instance, less than 20% of the size of the article on the failed alternative vote referendum, which is of course a topic which currently receives very little attention. This suggests that the quality of these articles, in Wikipedia terms, may be comparatively poor. However it also demonstrates that the quality of a Wikipedia article is not a perfect proxy for public attention: other factors drive the decisions of the community which edit these articles.

	Number of other Wikipedia articles linking in	Length (words)	Edits	Editors	Average Daily Page Views
Military Intervention in Libya	411	16,413	3,515	567	647
Alternative Vote referendum	281	14,974	1,286	220	152
Same Sex Couples Marriage Act	90	11,776	331	52	281
Tuition fees	35	9,417	472	97	278
Jobseeker's Allowance	99	3,429	355	96	237
Welfare Reform Act 2012	21	2,901	99	30	114
Police and Crime Commissioners	238	2,670	122	60	101
Free schools (England)	204	2,387	191	59	218
Department for Work and Pensions	350	1,964	301	94	165
Universal Credit	29	1,907	148	44	210
Fixed-term Parliaments Act 2011	39	846	81	24	103
Personal Independence Payment	28	773	24	9	31

Table 1: Size and edit history of different Wikipedia articles for the period January 2013 to October 2013

3.3 Mentions of DWP Policies on Social Media Platforms

In this section, we move on to explore the use of data from social networking sites such as Twitter and Facebook. This data can be useful for social research as a means of tracking both how much people are talking about specific topics of interest, and potentially what people are saying about them (i.e. they may fulfil both the public awareness and public opinion functions highlighted in section 2). As we highlighted in part 1, views expressed on social media platforms are not directly comparable to opinion polling, where questions are structured by those carrying out the survey and answers are offered anonymously. Rather, they provide a kind of insight into opinions which might be expressed in everyday conversation with friends (and hence perhaps have more similarity to a "focus group"). The large scale mining of such opinions has the potential to augment in a variety of ways our knowledge of both what topics the public are thinking about, and what their opinions are.

Here we look at two such possibilities. First, we use social media platforms to track trends in public attention to the policies of interest over a several month period. This allows us to make some speculative conclusions about how media events drive public attention, and which media stories truly created some kind of resonance in the social media sphere. Secondly, we use automatic sentiment detection techniques to try and map public opinion on the topic, based in particular on tweets on the subject.

3.3.1 Public Attention to UC and PIP on Social Media Platforms

We can get an idea of public attention to particular policies by measuring the number of users mentioning the policies on different social media platforms. Figures twelve and thirteen plot these mentions on Twitter and Facebook, the two most popular social media platforms (see section 1). We look at PIP, JSA and UC as usual. We also add terms for "removal of the spare room subsidy" as a means of comparison, which at the time the research took place was an important current policy initiative[47]. We look only at mentions of the full policies by name (e.g., "Universal Credit") to avoid measurement errors with the abbreviations used elsewhere (e.g., "UC" as University of California). We do allow for differences in capitalization and for "jobseeker's allowance" with and without the apostrophe as well as with or without a space between "job" and "seeker."

A number of conclusions are immediately apparent. Firstly, public attention is quite volatile: some days bringing high numbers of users mentioning the terms of interest, whilst others do not bring any at all. However this volatility is also different in different social networks. On Twitter, an individual search term can go from 1,500 users to 0 the following day, indicating that the "attention span" on this network is very short indeed, with short bursts of activity likely to be driven by specific media stories. On Facebook, by contrast, attention is a little more sustained. In the last month of sampling, for example, "removal of the spare room subsidy" was regularly mentioned by between 50 and 150 users each day. This indicates, perhaps,

[47] In the graphics below, absolute figures for this term are based on the aggregation of results from a number of search terms related to this policy change

Facebook's status as a more conversational network which has a somewhat longer attention span.

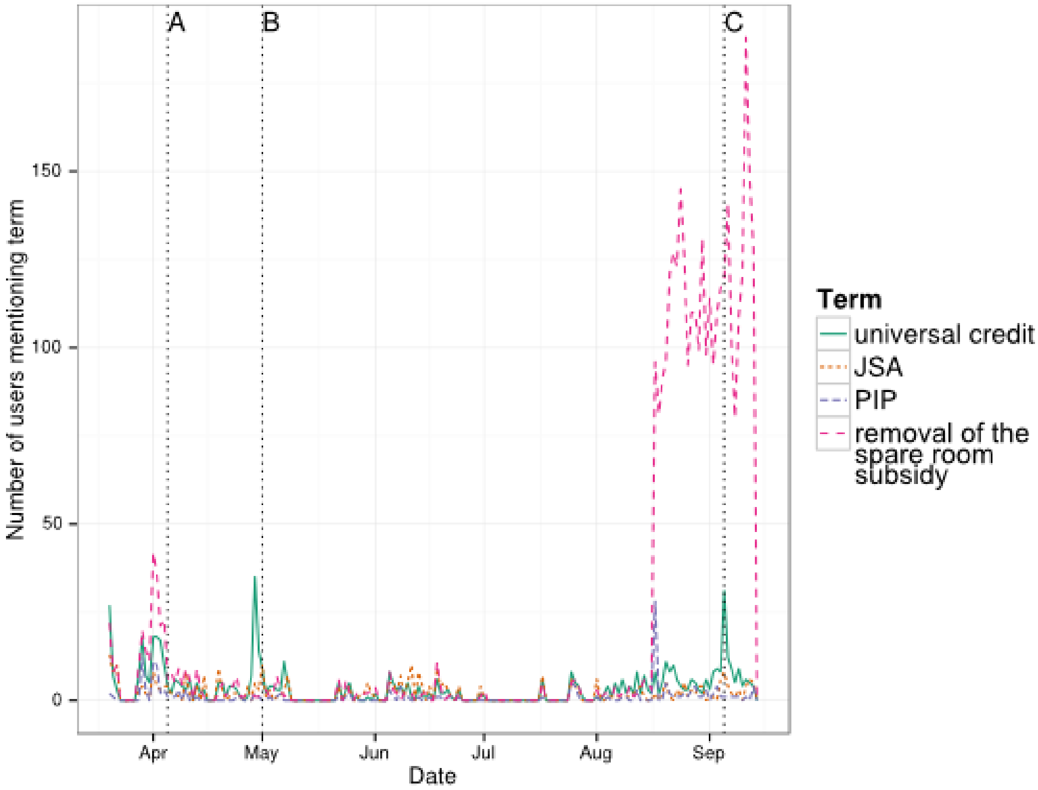


Figure 12: Number of Facebook users mentioning each policy per day. Vertical lines represent major news events

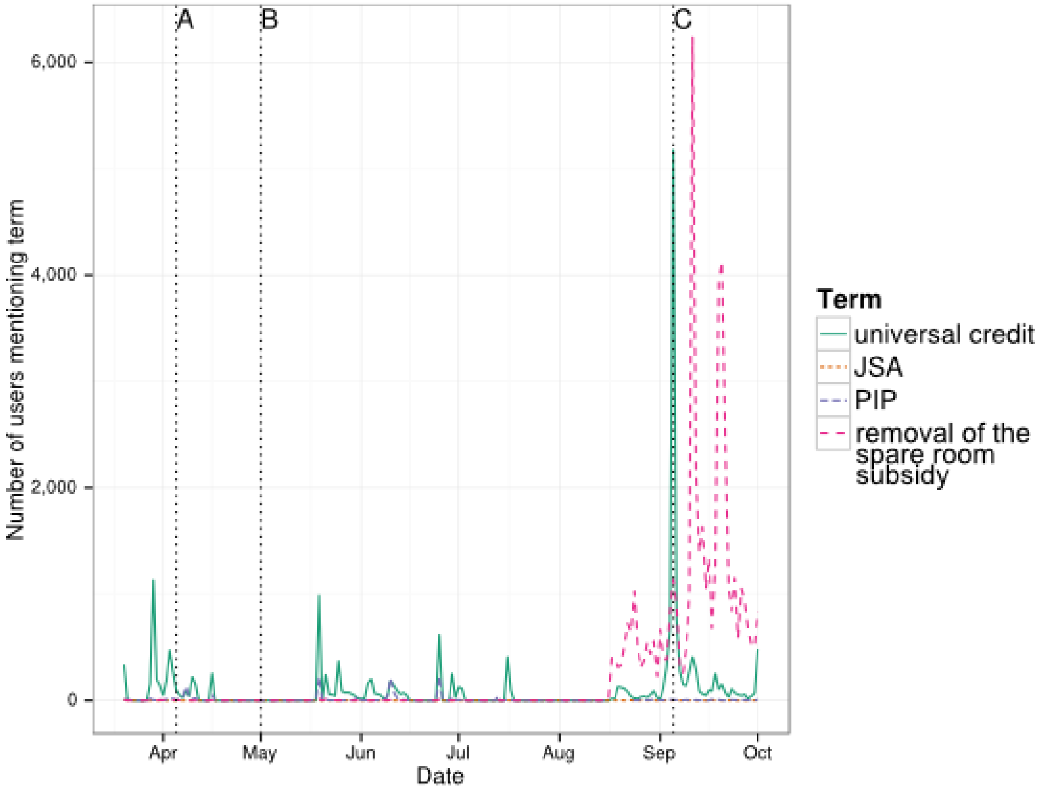


Figure 13: Number of Twitter users mentioning each policy per day. Two data collection issues led to missing data in May and August. Vertical lines represent major news events

Despite the volatility of daily measures, over the full time period two major conclusions can be drawn. Firstly, the overall numbers for all search terms are relatively low, with the exception of removal of the spare room subsidy which experiences a burst of activity towards the end of the sampling window. The numbers for our particular policies of interest are comparatively much lower than the number of Google searches and Wikipedia visits. This may indicate that these types of topics are ones that people prefer to search for information about in private, rather than discuss in public: though of course Universal Credit may start to see more discussion as implementation rolls out. Further research comparing to other policy issues would be needed to establish that definitively. Set against the overall picture of low use, however, there were some notable spikes in activity, particularly for Universal Credit towards the end of the sampling window.

The three news events related to Universal Credit are also highlighted in figures twelve and thirteen in the lines labeled A, B and C. We can see that event C at the beginning of September in particular drove a relatively large amount of Twitter reaction, and even a small Facebook spike. Hence significant media coverage of a particular policy does seem likely to provoke reaction in social media as well, and should be controlled for in subsequent analyses.

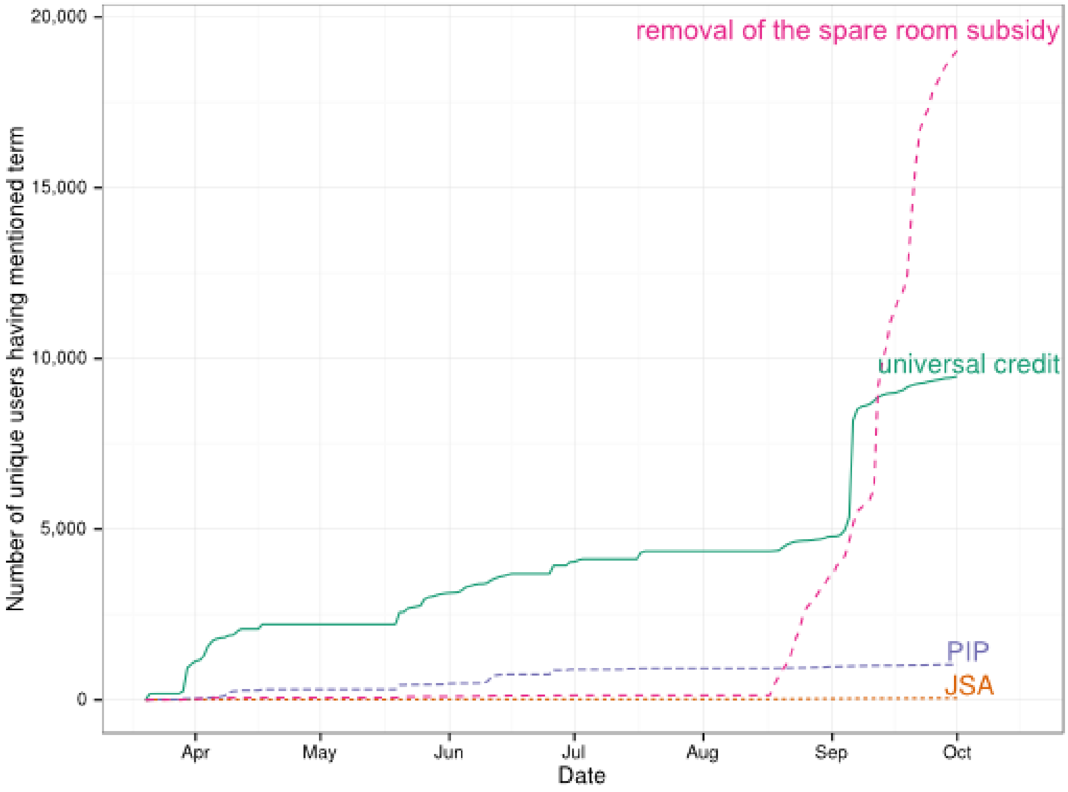


Figure 14: Cumulative Twitter users

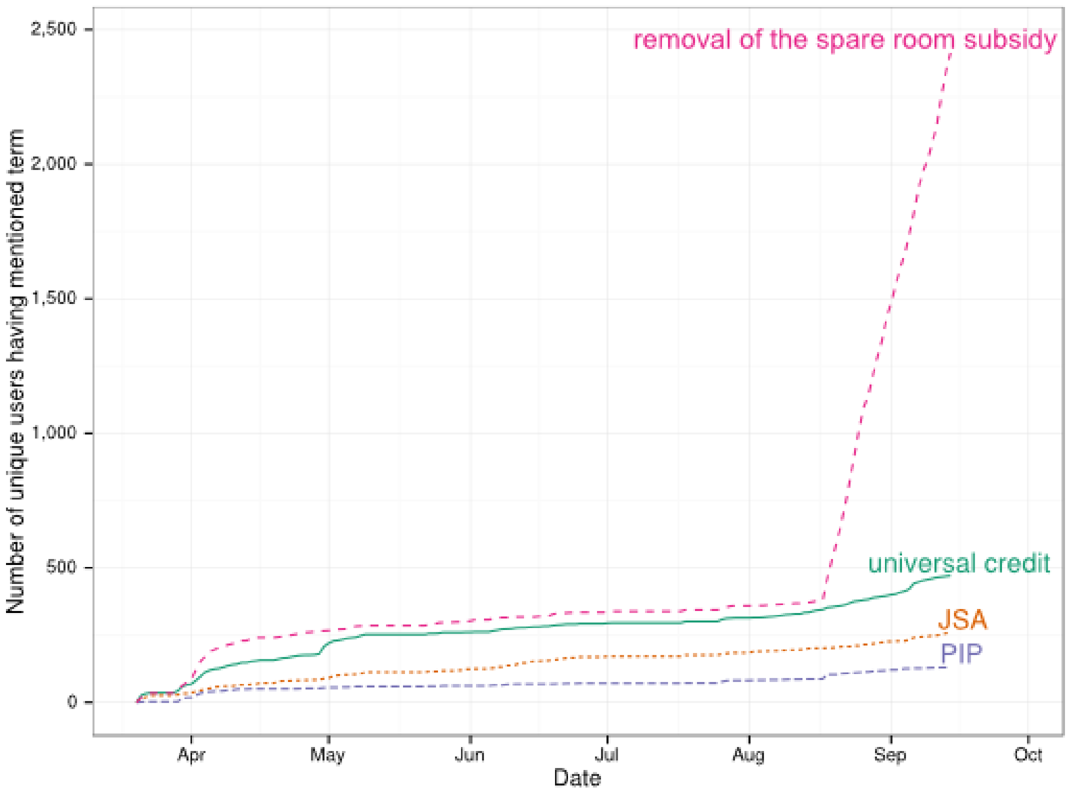


Figure 15: Cumulative Facebook users

It is also important to observe the extent to which people talking about a given policy have already mentioned it before. This provides an indication of the extent to which social network users are becoming more aware of the policy in question. We can tackle this question by looking at the cumulative number of users who have ever mentioned it on different social media platforms. This is something explored in figures fourteen and fifteen. On Facebook, we can see that all four search terms are conversations dominated by less than 500 users posting repeatedly, until mid-August when several thousand new voices are added to the conversation on removal of the spare room subsidy in a relatively short amount of time. On Twitter, we can see a roughly similar pattern: a slow accumulation of different users talking about Universal Credit up until September, whereupon around several thousand new ones are added.[48] These data again provide insight into exactly when the public at large starts to think about a given policy.

3.3.2 Social Media Data as an indicator of public opinion

Mentions on social media platforms may be used as an indicator of public opinion through the use of automatic “sentiment analysis”: a family of techniques which attempt to extract emotions expressed in texts. In this report we provide an example of such a technique: an automated “sentiment detection” program, developed by OII Associate Researcher Mike Thelwall and his research team called “SentiStrength”.[49] This program attempts to measure the strength of both positive and negative sentiment expressed in any short message on the basis of language use. It employs large dictionaries of keywords associated with positive and negative feelings, and assigns scores to texts mainly on the basis of the presence or absence of those keywords. Hence a keyword containing the word “hate”, for example, would register an increased score on the negative sentiment scale. SentiStrength also has a list of emoticons (mixtures of symbols and punctuation marks which are commonly used to express emotions, for example the “smiley face” formed from the symbols :-)) and a variety of extra rules to deal with negation and other linguistic issues. These positive and negative scores can then be summed together to produce a rough overall sentiment for any given tweet (though considering their different dimensions independently can also be helpful). As the developers of this tool note, the accuracy of this program is variable depending on context: one trial found around 60% accuracy for

[48] Difficulties with data collection mean that some of these new arrivals may actually have come in August
[49] See: http://sentistrength.wlv.ac.uk/

detecting positive emotions and a 73% accuracy for negative emotions.[50] This means that the results of any use of such sentiment analysis techniques have to be interpreted with care.

For the purposes of this example, we looked at around 20,000 original (non-retweet) tweets which were created during the sampling window which contained the phrase “Universal Credit”.

The results of such an experiment should be interpreted with caution. They represent the unsolicited opinions of a subset of Twitter users, hence cannot be simply translated into a score for public opinion on the program as a whole.

As we highlight in figure fourteen these come from approximately 10,000 individual users. Sentiment was, broadly speaking, found to be negative, with the majority of responses clustered in the range between 0 and -2 (on an overall scale which ran from -4 to 4).

Further research would be required to benchmark these expressed sentiments to other measures of public opinion over time to determine the extent to which they are truly representative of broader public opinion. The window of observation also remains relatively small, meaning that it is difficult to pick out trends in the data which might help confirm the overall validity of the approach. Nevertheless, what this section does highlight is the potential of sentiment analysis to provide a broad overview of a large set of comments on social media, in a way that would be very time consuming to do by hand. The next section contains further details and examples of automated sentiment analysis.

3.4 The DWP’s Own Social Media Presence

In this final section, we want to explore the potential social research uses of the DWP’s own social media presence (that is, all the accounts on all social networks which are maintained by various people working for the DWP itself). Of course, a primary use of many of these accounts is for communications and public relations. As we have said, we consider that application of social media to be beyond our scope here. However, these accounts may also be useful for more passive research purposes, as many people who wish to comment on DWP policies on social networks may direct their comments directly at DWP accounts. Exploring the nature of such comments may provide hence another useful window on public opinion. Furthermore, as we highlighted above, one key problem of opinion polling lies in finding “sub-populations” of interest, for example jobseekers in a specific

[50] Thelwall, M., Buckley, K., Paltoglou, G. Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12), 2544–2558.

area. The extent of the DWP's Twitter presence presents an opportunity to fill in this gap somewhat. By collecting tweets directed by people towards specific Jobcentres, we can start to form a picture of what people using these Jobcentres are thinking about, and (perhaps) how they evaluate the way these centres perform.

In order to start to explore this potential, over a two week period we collected all tweets mentioning the name of any of the DWP's active Jobcentre accounts (which include both individual Jobcentres and some regional accounts which cover wider areas). We exclude tweets created by one of the Jobcentres themselves, or simple "retweets" of messages created by these centres. Within the period of observation, about 1,000 tweets of the 20,000 in total collected fulfilled these criteria.

We employed the same sentiment detection process described in section 3.3.2. As we argue above, automatic content analysis is not yet really useful for analyzing specific topics which people are complaining about. However, where it can be useful is in terms of performing an initial filter when searching for certain types of tweets. For example, if we are interested in people expressing feedback about a particular Jobcentre or service, we could use sentiment analysis to collect all tweets with a low score on the negative sentiment scale, on the basis that it is difficult to express criticism without using any critical words (bearing in mind, however, the points about irony and sarcasm). Using this method of filtering, seven tweets were identified, detailed in table three. Of these tweets, only the first is really misclassified (in that it does not express negative sentiment), with some of the rest providing quite clear feedback about people's experience of specific Jobcentres (though it must be noted that the requirements of anonymisation mean that interpretation of some of the text as presented below is more difficult, as they have been stripped of all contextual information).

ID	Tweet Text	Negative Sentiment Score. Scale from 1 (lowest) to 5 (highest)
1	@TWNAME1 Are you over 50 and interested in becoming your own boss? Try our training course in PlaceA	3

	http://t.co/linkname	
2	@TWNAME2 these companies are determined to tip families over the edge. It is a disgrace that the government dont do anything	4
3	@TWNAME3 tried calling today...spent so long on hold. In the end phone went dead	3
4	@TWNAME4 I was told to call, to change time I would be seen but the phone doesn't get answered and then goes dead. Pls advise urgently	3
5	@TWNAME5 does that include your colleagues down the road who've been offered redundancy?!	3
6	@TWNAME6 ask them to stop offering jobs to kids and give us jobless grown ups a chance as age discrimination is against the law.	3
7	@TWNAME7 awful website, you can't just search for PlaceA, Region - Place great description!! .... not	4

Table 2: Directed feedback to DWP Jobcentres. Names and places have been removed.

This brief trial highlights three things. Firstly, it is possible to collect quite detailed feedback about individual Jobcentres, which is apparently delivered by people with specific experiences of using the centre shortly after their experiences took place. Some of it provides a potentially useful insight into the way working practice could be changed. For example, tweet number 7 would like for the website listing jobs in the Jobcentre of interest to offer an extra filter box which allows them to search for jobs just from a specific area. Secondly, however, it is not possible to automatically extract the precise meaning of the text. While we can detect with relatively good accuracy when someone is complaining, human intelligence is required to know *what* they are complaining about. This implies that significant manual work would be required to get full benefit out of any mechanism which tried to capture feedback automatically.

Finally, however, it is worth noting that, even in the context of a short two week trial, the amount of data collected is small. Any experiment to, for example, evaluate the performance of Jobcentres based on Twitter feedback would have to take place over a relatively long time period, or be unified with other data from other social media sites; however, even then it would be wise to interpret the results with caution. Furthermore, we observed no tweets about the specific topics of PIP and UC (apart from those created by official DWP accounts). On the basis of our observations above about the relatively low total number of mentions of these specific policies, this is not surprising. However it does mean that using tweets directed at DWP accounts as a means of evaluating specific policy proposals such as UC and PIP is not really feasible at this stage. This might be something that would increase as the policies themselves are rolled out further.

4 Conclusions and Further Work

In this section, we set out our conclusions for the project, and recommend directions in which the Department for Work and Pensions could take further work. The main aim of this report, and the research project on which it is based, was to explore how social media data could be used to help support policy research and analysis at the DWP. In particular, the aim was to explore the extent to which such data could act as a reliable source of public opinion in respect of two key Welfare Reform policies: Universal Credit and Personal Independence Payment.

In general terms, we think that a broad consideration of DWP policies across a range of social platforms has the potential to inform research and policymaking in a variety of ways. Some of these have been highlighted in the empirical section. We have seen how social data makes it possible to explore the extent to which people are thinking about a particular topic, and where they go for information on that topic. In this regard, the DWP should consider both the position of its own website in the social landscape, and the accuracy of any other competing websites. We have shown how public attention responds to specific media events, and that part of an effective information strategy involves knowing when the public will start looking for information on a particular subject. In terms of Universal Credit, we can see that searching for the topic has been going on since it was first announced in 2010. However, the Wikipedia page remains a relatively poor source of information on the topic (compared to other similar Wikipedia pages).

We have also seen how social media can be useful for tracking expressed reactions to a particular policy. We have highlighted that these reactions cannot be easily generalized to broader “public opinion”, and that methods for analyzing their sentiment remain a work in progress. Nevertheless social media do provide an impression of how much a topic is being discussed, and the total spread of its awareness around people using social media. We have also shown, in the case of UC and PIP, that social media reaction is much smaller than average numbers of people searching for information on a particular topic.

There are, however, a variety of opportunities which we were unable to fit into the relatively short context of this report, but that could be explored in further work. Most obviously, a variety of further social media platforms could provide useful insight, from LinkedIn to the government’s petitions site (where there are over 250 petitions on Universal Credit), to more specific forums such as Mumsnet and Money Saving Expert, which present real potential in terms of contacting specific groups of individuals. The overall data collection period for the trial was also quite small, especially in the context of two policies which are only just starting to penetrate into public consciousness. Further work taking in a much longer time period would be useful in terms of further validating measures of public opinion.

Finally, we argue that further work ought to benchmark social media data to other existing sources of information. Server logs of own web pages would be extremely valuable in this context: we recommend the Department consider how traffic statistics change on individual pages as a useful further measure of what the public is currently interested in, and benchmarking this against the results from Google Trends. Existing opinion polls and survey research would also be another useful resource in this context: it would be interesting to see the extent to which they match results of sentiment analysis conducted on the social web. Overall, we have emphasised the need to remain sceptical to some of the tools being developed by social media analysis. They are in their early stages, and are unlikely to replace more traditional research methods such as the sample survey in the near future. Nevertheless the possibilities for social media data are considerable, and we strongly recommend the DWP continues to work in this area.

5 References

Acquisti, A. and Gross, R. (2006). "Imagined communities: Awareness, information sharing, and privacy on the Facebook". *Privacy Enhancing Technologies, Lecture Notes in Computer Science*, 4258, 36–58

Bartlett, J., Birdwell, J., and Littler, M. (2011). *The New Face of Digital Populism*. London: DEMOS.

Blank, G. (2013). "Who creates content?" *Information, Communication and Society*, 11, 2, 590-612.

boyd, d and Ellison, N. (2007). Social Network Sites: Definition, History and Scholarship. *Journal of Computer-Mediated Communication*, 13, 1, 210-230

Bruns, A. (2008) The Active Audience: Transforming Journalism from Gatekeeping to Gatewatching. In *Making Online News: The Ethnography of New Media Production*. Eds. Chris Paterson and David Domingo. New York: Peter Lang

Choi, H., and Varian, H. (2012). "Predicting the Present with Google Trends". *Economic Record*, 88, 2-9.

Dutton, W., and Blank, G. (2011). “Oxford Internet Survey 2011: Next Generation Internet Users”, 34. Available from: http://oxis.oii.ox.ac.uk/sites/oxis.oii.ox.ac.uk/files/content/files/publications/oxis2011_report.pdf

Dutton, W., and Blank, G. (2013). “OxIS 2013 Report: Cultures of the Internet”. Available from: http://oxis.oii.ox.ac.uk/sites/oxis.oii.ox.ac.uk/files/content/files/publications/OxIS_2013.pdf

Feldman, R. (2013). "Techniques and Applications for Sentiment Analysis". *Communications of the ACM*, 56(4), 82-89

Furnas, A., and Gaffney, D. (2012). “Statistical Probability That Mitt Romney's New Twitter Followers Are Just Normal Users: 0%”. *The Atlantic*. Available from: http://www.theatlantic.com/technology/archive/2012/07/statistical-probability-that-mitt-romneys-new-twitter-followers-are-just-normal-users-0/260539/

Ginsberg, J., Mohebbi, M., Patel, R., Brammer, L., Smolinski, M., and Brilliant, L. (2009). "Detecting influenza epidemics using search engine query data". *Nature*, 457, 1012-1014

Golbeck, J., and Hansen, D. (2013). "A method for computing political preference among Twitter followers". *Social Networks*, 36, 177-184.

Gonzalez-Bailon, S., and Wang, N. (2013). "The Bridges and Brokers of Global Campaigns in the Context of Social Media". Available from: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2268165

Graham, M., Hale, S. A., and Gaffney, D. (2013). "Where in the World are You? Geolocation and Language Identification in Twitter". *Professional Geographer*, forthcoming.

Hargittai, E., and Walejko, G. (2008). "The Participation Divide". *Information, Communication and Society*, 11, 2, 239-256.

Kosinski, M., Stillwell, D., Graepel, T. (2013). "Private traits and attributes are predictable from digital records of human behavior". *PNAS*, 110(15), 5802-5805

Margetts, H., and Hale, S. (2010) "User Experiments. Phase 1: Life Events Report". *Study on user expectations of a life events approach for designing e-government services*, European Commission, 79-96.

Mestyán, M., Yasseri, T., and Kertész, J. (2013). "Early Prediction of Movie Box Office Success Based on Wikipedia Activity Big Data". *PLoS One*, 8(8)

Newman, M., Groom, C., Handelman, L., and Pennebaker, J. (2008). "Gender Differences in Language Use". *Discourse Processes*, 45(3), 211-236

Office of National Statistics. (2012). “Internet Access - Households and Individuals, 2012 part 2". Available from: http://www.ons.gov.uk/ons/rel/rdit2/internet-access---households-and-individuals/2012-part-2/stb-ia-2012part2.html

Office for National Statistics. (2012) “Workless Households for Regions across the UK, 2012”. Available from: http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77-322248

Office of National Statistics. (2013). "Internet Access – Households and Individuals, 2013”, Available from: http://www.ons.gov.uk/ons/dcp171778_322713.pdf

Ortega, F., Gonzalez-Barahona, J., and Robles, G.(2008). "On the inequality of contributions to Wikipedia". *HICSS '08 Proceedings of the 41st Annual Hawaii International Conference on System Sciences*, 304

Preis, T., Moat, H., and Stanley, H. (2013). "Quantifying Trading Behavior in Financial Markets Using Google Trends". *Scientific Reports*, 3, 1684

Ripberger, J. (2011). "Capturing Curiosity: Using Internet Search Trends to Measure Public Attentiveness". *Policy Studies Journal*, 39, 2, 239-259.

Samoilenko, A., and Yasseri, T., (2014) The distorted mirror of Wikipedia: a quantitative analysis of Wikipedia coverage of academics, *EPJ Data Science*.

Sullivan, D. 2013. "Google Still World’s Most Popular Search Engine By Far, But Share Of Unique Searchers Dips Slightly". *Search Engine Land*. Available from: http://searchengineland.com/google-worlds-most-popular-search-engine-148089

Sweeney, L. (2002). "*k*-anonymity: a model for protecting privacy". *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, 10(5), 557-570.

Thelwall, M., Buckley, K., Paltoglou, G. Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61, 12, 2544–2558.

White, D., and Le Cornu, A. (2011). "Visitors and Residents: A new typology for online engagement". *First Monday*, 16(9).

Wilkinson, D., and Huberman, B. (2007). Cooperation and quality in Wikipedia. *WikiSym '07 Proceedings of the 2007 international symposium on Wikis.*

Wright, F., White, D., Hirst, T., and Cann, A. (2014). "Visitors and Residents: mapping students attitudes to academic use of social networks". *Learning Media and Technology*, 39, 1, 126-141.

Yasseri, T., Sumi, R., and Kertész, J. (2012). "Circadian Patterns of Wikipedia Editorial Activity: A Demographic Analysis". *PLoS One*, 7(1).

6 Annex

7 This is a full list of the keywords which were tracked across social media platforms during the creation of the report. Hash signs were included at the front of certain keywords indicated below to facilitate monitoring on Twitter.

8

Personal Independence Payment
Personal Independence Payments
(#)PIP
Disability Living Allowance
DLA reform
ATOS
Capita assessment
face to face consultation
PIP assessment
PIP rules
PIP rates
PIP claim
PIP assessment
PIP consultation
PIP mobility
PIP eligibility
PIP criteria
Sickness benefit
"on the sick"
"on Disability"
(#)UC
Universal Credit
Tax credit
Benefits
(#)JSA
(#)DSS
Department for Social Security
"on benefits"
"on the social"
Unemployment Benefit
Welfare reform
Benefit reform
Benefit Changes
Welfare changes
Benefit Cuts
"The Dole"
"The Social"
Under occupancy Charge
Removal of the Spare Room Subsidy
Benefit Cap
Benefit Cuts
Child Tax Credits
Employment and Support Allowance
"Jobseeker's Allowance"
"Jobseekers Allowance"

"Job seekers Allowance"
"Job seeker's Allowance"
Working Tax Credits

9