Chapter 9

Rethinking Democracy with Social Media

HELEN MARGETTS

SOCIAL MEDIA are blamed for almost everything that is wrong with democracy. They are held responsible for pollution of the democratic environment through fake news, junk science, computational propaganda and aggressive micro-targeting and political advertising. They are accused of creating political filter bubbles, where citizens exist in ever narrower 'echo chambers' of personalised news and connections with like-minded people, which mean that they are exposed only to similar ideological viewpoints, feeding their own opinions back to themselves and creating a '*Daily Me*' news environment. In turn, these phenomena have been implicated in the rise of populism, political polarisation, waves of hate against women and minorities, far-right extremism and radicalisation, post-truth, political chaos, the end of democracy and ultimately, the death of democracy.[1] Discussion of social media's role in democracy sounds like a premature lament for a sick patient, without investigating the prognosis.

Yet, actually we know rather little about the relationship between social media and democracy. In their ten years of existence, social media have injected volatility and instability into political systems, bringing a continual cast of unpredictable events. They have challenged normative models of democracy—by which we might understand the macro-level shifts at work—seeming to make possible the highest hopes and worst fears of republicanism and pluralism. They have transformed the ecology of interest groups and mobilisations which challenge elites and ruling institutions, bringing regulatory decay and policy sclerosis. They create undercurrents of political life that burst to the surface in seemingly random ways, making fools of opinion polls. But although the platforms themselves generate new sources of real-time transactional data that might be used to understand this changed environment, most of this data is proprietary and inaccessible to researchers, meaning that the revolution in big data and data science has passed by democracy research. This

chapter looks at the available evidence regarding the effect of social media on democracy, for which—as for Mark Twain—the report of death may be an exaggeration.[2]

What do we know? The value of small things

Social media—digital platforms which allow the creation, location and exchange of content—are entwined with every democratic institution and the daily lives of citizens, having reached incredible levels of penetration. Worldwide, Facebook has 2 billion users, YouTube has 1.5 billion, WhatsApp 1.2 billion, Instagram 700 million, Twitter 328 million and the Chinese WeChat 889 million; nearly three quarters (73 per cent) of US adults use YouTube, while 68 per cent use Facebook.[3] In the UK, 66 per cent of the population are active social media users, with 57 per cent using social media on mobile (these groups overlap, with many using both).[4] Among younger age groups, usage of at least one social media platform is nearly ubiquitous, with 60 per cent of teenage smartphone users using Snapchat alone. When deciding whether to vote, to support, to campaign, to demonstrate, to complain, digital media are with us at every step. They shape our information environment through search engines and social information about what other people are doing, and extend our social networks by creating hundreds or thousands of 'weak ties', particularly for users of social media platforms such as Facebook or Instagram. So, these platforms have transformed the costs and benefits of every kind of political participation. But the key difference that social media have brought to the democratic landscape is a raft of new activities which are characterised by being really small, extending below the bottom rung of the ladder of participation. These activities stretch from small acts such as signing a petition, through voting, attending a political meeting and donating money to a political cause, right up to political violence or armed struggle. Following, liking, tweeting, retweeting, sharing text or images relating to a political issue, or signing up to a digital campaign, are tiny acts of political participation that have no equivalent in the pre-social media age (there is no precedent, for example, for reading President Trump's tweets). Even tweeting about a demonstration or political event that you have not attended is an act, because you have sent a tiny signal of viability to anyone looking at your tweet (or the +1 on a 'like'), and made it that bit more likely that they will act. These tiny acts are all around the democratic environment. They have enabled ordinary people across the world with no more resources

than a mobile phone to challenge injustice, fight for policy or regime change, and shed light on corruption and inefficiency in public life. Until only twenty years ago, there was no way to participate in politics without joining a political party or organised interest group, attending meetings and knocking on doors. For many people, these costs in terms of time, effort or resources were too great, and politics has often been the province of an activist elite. Now, tiny acts are drawing new people into politics, particularly young people, whose absence political commentators have been bemoaning for years. For example, it is hard to imagine in a pre-social media era that US school children would demonstrate, under the hashtag #Enough, and walk out of school to campaign for gun control, as they did in 2018 in the wake of yet another school shooting.

Scaling up—or not

Taken individually, tiny acts of participation seem insignificant, and indeed for many years were dismissed as mere ‘slacktivism’ or ‘clicktivism’, denigrated as low-cost political acts that have minimal effect. The US political commentator, Malcolm Gladwell, published a widely cited *New Yorker* article[5] in 2010, arguing that small-scale actions and weak ties facilitated by social networking platforms could never give rise to political mobilisation on the scale of the civil rights movement, which provoked some controversy given the demonstrations, protests and even revolutions of the so-called Arab Spring which followed so soon afterwards. Tiny acts can and do scale up to large-scale mobilisations and campaigns for policy change that have brought major shocks and surprises to political regimes all over the world. Petitions to re-run the UK referendum regarding membership of the European Union (in 2016), or to block Donald Trump from a state visit to the UK (in 2017) immediately shot up to millions of signatures. There can be hardly a country in the world that has not experienced widespread demonstration, protest or campaigning for political change taking place on social media—demonstrations against the financial crisis and austerity in the US and UK in 2009 being examples of such. These mobilisations have challenged and even brought regime change in authoritarian states, as well as a whole host of policy changes in liberal democracies, from justice for victims of police brutality to the end of controversial health reforms. A few of these mobilisations have developed into new political parties across the ideological spectrum, from the Spanish Podemos

to the Italian Five-Star Movement to the German far-right Pegida, all highly disruptive forces in the political systems in which they have emerged. This scaling up of tiny acts appear also to play a part within conventional political events—such as election campaigns—by building up into waves of support for unconventional candidates such as Donald Trump (elected as US President in 2017), or Jeremy Corbyn (elected as leader of the UK Labour party in 2015 and again in 2016).

So, tiny acts of participation can scale up dramatically and rapidly. But they almost always don't. It may seem from the news media that it is relatively easy to get hundreds of thousands of people out in the streets or into a square, given the frequency with which such events appear to occur. But there is an obvious selection bias in favour of the successful mobilisations, which are reported on TV screens and circulate on social media, compared with all the failed mobilisations that we never see. So, for example, in the UK, over 99 per cent of petitions to the government fail to get the 10,000 signatures required for an official response and only 0.1 per cent attain the 100,000 required for a parliamentary debate (the same figure is 0.7 per cent in the US). This picture is replicated over and again across social media platforms from YouTube to Facebook, and across countries. There is no normal distribution of mobilisations; rather, a distribution that is more like that of earthquakes—a small number of extreme events and a huge number of insignificant ones. It is hard to predict (just as it is for earthquakes) which will succeed and which will fail. The shift towards new forms of mobilisation based on tiny acts of digital participation has brought an era of what we labelled in our book of the same name *Political Turbulence,*[6] where politics becomes harder and harder to predict.

Political turbulence means a challenge to two stabilising elements of democracy—political identity and institutions. Rather than identifying with issues, forming collective identity and then acting to support the interests of that identity (or voting for a political party that supports it), in a social media world, people act first, and think about it, or identify with others later, if at all. And even when a mobilisation does succeed in getting a million people on the street, or signing a petition, or even a revolution of sorts as in Egypt, the very fact that it is possible to get there without the normal organisational trappings (such as nascent parties or leaders) of a movement or revolutions means that it will usually be unsustainable. At the same time, turbulence threatens democratic institutions of all kinds, by showing that political figures from outside the mainstream—such as Jeremy Corbyn and

Donald Trump—can win, or at least garner huge levels of support in surprising ways in a short space of time. Will the Labour party ever recover from Corbyn, the Republican party from Trump? Social media seem to inject instability into traditional democratic institutions.

Losing control of democracy

While traditional institutional actors are struggling in this new democratic landscape, new ones, of vital importance, have emerged. That is, when citizens are deciding to undertake one of these new tiny acts of political participation, they do so on digital platforms—and in particular, social media platforms. And these platforms shape crucially the information environment within which these decisions are made. They do so by exerting social influence, in the form of information about what other people are doing. In an earlier era, when we signed a petition on a street or expressed our support for some campaign, we knew little about how many other people supported the cause. On social media, on digital activism platforms, we know straight away how many people like, follow, share, view or discuss an issue or news item, and how close a campaign is to reaching its target. And decades of social science show that information like this is a crucial influence on how much we ourselves support something. For example, an experiment with music platforms has shown that if people are told that songs are popular, they are more likely to rate the songs highly themselves in comparison with being told they are unpopular, or being given no information at all, while we are more likely to contribute to public goods by recycling or voting or give to charitable causes if we know that other people are doing so.[7] This kind of influence has been shown to operate in the same way on digital platforms. For example, trending information, showing the most popular petitions, has been shown to concentrate petition signing in those petitions at the expense of the petitions not making it to the trending box.[8]

These chains of social influence, by which individuals undertake acts of participation and thereby send signals of viability to others, are the explanation for how small acts can lead to large-scale mobilisation or waves of support. This can start primarily on one platform, as with the so-called 'Twitter revolution' of Tunisia or 'Facebook revolution' in Egypt in 2011, or the largest demonstration in Romanian history in the autumn of 2017, with widespread protests against corruption in a newly

elected government in Romania attempting to pardon themselves for past crimes, which appears to have been coordinated via the messaging app Slack. More recently, chains of influence operate across multiple platforms, particularly where young people are involved (with most teenagers in the US and UK being active on up to five social media platforms). In all these events, social media sent signals of viability through likes, shares, follows, views and so on, leading people to consider that mobilisations might succeed and be worth joining. That makes the digital media platforms themselves crucially important democratic actors, particularly the most popular giants of the social media world: Google (which also owns YouTube) and Facebook (which owns Instagram and WhatsApp). The design of their platforms shapes political participation in terms of what information people encounter, how people share information, and what they know about what other people are doing.

Furthermore, while a rise in political mobilisation and activism might be considered generally a positive development for democracy, the same mechanisms—tiny acts and chains of social influence—can also lead to anti-democratic phenomena. Just as social media platforms allow small acts of participation to become widely accessible and to (sometimes) scale up, they allow acts of misinformation, hate speech, abuse, threats, extremism, radicalisation and even terrorist influence to follow the same process. For this reason, social media platforms are implicated in a number of pathologies of contemporary democracy, including:

- *Echo chambers*, in which people are surrounded in online social networks by like-minded people and opinions that reinforce their own belief systems (in the same way that acoustic echo chambers use hollow enclosures to produce reverberated sounds);
- *Fake news*, where distorted or false versions of events are widely disseminated either for the purposes of disruption or for financial gain;
- Highly targeted *political advertising* to the extent that it is personalised, based on personal data;

- *Computational propaganda*, involving automated social media accounts (bots) which mimic real people through the dissemination of information or fake news across a range of platforms and networks, with the intention of manipulating opinion;
- *Hate speech*, where online abuse or threats are directed at individuals or groups on the basis of attributes such as race, religion, ethnic origin, sexual orientation, disability, or gender.

Critics of social media's role in democracy view these phenomena as operating in tandem. That is, automated bots are used to spread disruptive rumours in the form of fake news, which are spread around and reinforce the ideological position of echo chambers, whose inhabitants are vulnerable to these distorted versions of events and regard them as the truth. They are also used to disseminate hate speech to magnify its effect, surrounding victims with tirades of abuse and harassment from both human and automated 'trolls'. In this way, echo chambers lead to polarisation, dragging those in the middle ground towards more extreme opinions on either side of a duality, such as whether or not to remain in the European Union or whether to vote for Donald Trump. The view of these phenomena acting in concert rose to the surface of political commentary in March 2018 with the Cambridge Analytica affair (described in Martin Moore's chapter), where Facebook data generated by an application purportedly for academic research purposes was revealed to have been sold by the developer (thereby breaking Facebook's terms and conditions) to the political campaigning organisation Cambridge Analytica. The company then used this data for highly targeted advertising in the campaign to elect Donald Trump and (allegedly) in the UK referendum on EU membership. While none of the other protagonists in this affair—the Cambridge academic who developed the application, Cambridge Analytica, or the whistleblower Chris Wylie, who worked for the company—come out well, Facebook in some ways received the most opprobrium. For adherents to this view, the acoustics of social media, orchestrated by firms like Facebook, are implicated in the waves of political populism and even extremism that have swept across the United States and many European countries.[9] Although all of these phenomena—and in particular, the effect that they have—is contestable and under-researched, as discussed below, the perceptions that social media are somehow implicated in the downfall of democracy are real, and the journalist Carole Cadwalladr who exposed the Cambridge

Analytica scandal believes strongly that social media platforms are leading us into a '9/11 of democracy'.[10]

Democratic grief

Political turbulence, the rising influence of social media platforms, and the loss of control for traditional institutions is traumatic for democrats. As noted in the introduction, many approach it as a death, with something akin to grief. It is possible to identify within the debate responses analogous to the famous five stages of grief: denial, bargaining, anger, depression and acceptance.[11]

Denial: politics as usual

The first response to the relationship between social media and politics has been denial. The way this view goes is that there is nothing new here, and that digital technologies in general are a neutral tool that (sometimes) make things work better (as Jane Austen might put it).[12] Although computer systems entered the machinery of government as early as the 1950s, when they first started to be used for large-scale administrative processing in the UK and US, they have been largely ignored in the media view of government outside the trade press. Before the 2000s, visual images of politics or political movies never contained technology—always venerable buildings, talking heads, pens, and desks. Any consideration of technological change still has no part in politics courses. Part of this resistance comes from the 'politics as pain' principle—particularly in British politics—that insists that politics that is automated or digitised in some way is not real politics. It should involve a long boring meeting or knocking on doors in bad weather or cut into the evenings, the view behind the dismissal of social media's role in politics as mere 'clicktivism' or 'slacktivism'.[13] It is interesting that this view manages to prevail, even while social media are being blamed for their massively pernicious influence.

Bargaining: the internet will make us free

Others take the second stage of grief—bargaining—to extreme. Here is the view that technology in the form of the internet is going to transform our political system and solve the traditional dilemmas of politics. We can live in a hyper-modernist world of direct democracy, where the bureaucratic state can

disintegrate. But this can happen only if we preserve the internet as an icon of freedom, unconstrained by governance or regulation, following the original cyberactivist John Perry Barlow[14] in his Declaration of the Independence of Cyberspace, which proclaimed:

> Governments of the Industrial World, you weary giants of flesh and steel, I come from Cyberspace, the new home of Mind. On behalf of the future, I ask you of the past to leave us alone. You are not welcome among us. You have no sovereignty where we gather. We have no elected government, nor are we likely to have one ... I declare the global social space we are building to be naturally independent of the tyrannies you seek to impose on us. You have no moral right to rule us nor do you possess any methods of enforcement we have true reason to fear.

In this world, political life will be reinvented, a 'global social space we are building to be naturally independent of the tyrannies you seek to impose on us' as long as we are true to this cyber-libertarian dream. Even in a political world dominated by social media platforms run by huge internet corporations such as Google and Facebook (that might seem to have replaced the 'giants of steel'), adherents to this faith dream on, and for them any censorship or regulation of the internet or social media is opposed vehemently. Even hate speech and micro-targeted personalised advertising must all be allowed to continue unchecked.[15]

Anger: shoot the messenger!

This is currently the dominant view, where technology—particularly the internet, and over the last decade, social media—is to blame for everything bad in democracy, from the rise of radicalisation and the extreme right (or left), to the election of Donald Trump as US President in 2016. This is a special case of a more general tendency to blame the internet for everything, from plane crashes to teenage suicides. Here, social media is responsible for most pathologies of modern election campaigns, through the creation of echo chambers or filter bubbles of like-minded people, where citizens receive constant reinforcement of their own views, which somehow leads to political polarisation. These

bubbles are reinforced in a number of ways, including the use of fake news, where third parties (non-media organisations) create completely untrue stories and tempt people to read them, thereby making money of advertisements alongside; political bots (robotic social media accounts that give the impression that a political campaign has more supporters than it actually has); or attempts generally to try to disrupt an opponent's campaign. The other accusation made against social media is that of hate speech (through the phenomenon of trolling against public figures, particularly women or people of the Muslim faith) and, more generally, a degeneration of civic discourse and a further polarisation of political discussion, as people on one side of a debate (such as whether to leave the European Union) react strongly to abusive language on the other.

Depression: post-truth

Depression is the next stage, where social media have led us to a post-truth world, where we cannot distinguish real news from fake news, and 'objective facts are less influential in shaping public opinion than appeals to emotion and personal belief'.[16] Under this view, the internet corporations will continue steadily in their inexorable rise, taking over some roles from the state and turning citizens into ad-clicking data providers. Meanwhile, any genuinely 'social' media will be strangled, choking on their own hate, dominated by computational propaganda from massive political entities (such as the Russian regime) so that genuine political movements as characterised by the Arab Spring could never happen again. For some, this view is bound up in a more general rejection of technology's role in society, for example, in automating jobs, and a desire for a return to the past, which manifested itself so powerfully among Trump supporters in 2016, with their MAGA (Make America Great Again) hats, or in the Leave campaign in the UK referendum, with their (ironically, post-truth) red bus.

Acceptance: moving on from myths

Reaching the final stage of grief would be to accept that digital media platforms are part of our democratic system, the political weather, and that political systems must accommodate the change, through a process of institutional catch-up. Most social media platforms did not exist ten years ago, and they have been at the heart of our political systems for far less than that, so it is understandable

that political institutions have failed to adjust, and the new institutions of democracy—social media corporations—have proceeded unchecked and unregulated, particularly given the power of the original cyber-utopian dream. In many ways, the myths of the other stages of grief—the denial of the importance of the internet, the denigration of social media activism as slacktivism and clicktivism (denial), the insistence that we can never regulate internet-related platforms (bargaining), the belief that social media are to blame for everything (anger) and the hopelessness of taking back any sort of control (depression), all work against the possibility of integrating and institutionalising social media platforms into democratic life.

The first stage in acceptance therefore would be to tackle some of these myths with research, to understand the scale and scope of the democratic pathologies outlined above. Such a task is not as straightforward as it might seem; ironically, because these platforms are based entirely on data, it is extremely difficult obtain the data needed to understand democratic processes, institutions and behaviour. While Twitter data is relatively open (hence the disproportionately high levels of scholarly attention it receives), WhatsApp is encrypted, Snapchat data is deleted as soon as it is read and so on. Most social media data from the more popular platforms such as Facebook and Instagram (which is owned by Facebook) is proprietary and closely guarded, especially in the wake of the Cambridge Analytica affair of 2018. Incidents like this, as with AOL's release of 1 million search records for research purposes in 2006, which were de-anonymised in a few minutes of release, have done much to stifle research into social media's actual (as opposed to gloomily hypothesised) impact upon democracy.

Examining the evidence for the death of democracy

One key research question is whether echo chambers and filter bubbles really exist to any greater extent than in non-digital settings (such as just reading the *Daily Mail* or only watching Fox News). After all, people have always sought out like-minded people and society has been structured around families. It is undoubtedly the case that social media firms can shape the kind of news environment and social networks that form on their platforms, for example through whether they provide trending information (Twitter always does this and Facebook did for a several years but stopped doing so in

2018), or by showing certain elements of the news feed. (Facebook took the decision at the beginning of 2018 to prioritise interactions with friends and family, rather than passive consumption of news, arguing that this was more conducive to well-being.) The most ardent believer in the echo chamber phenomenon, the US legal scholar Cass Sunstein, has moved from his pessimistic 2001 and 2007 predictions for the use of search engines for news consumption—made before use of social media became widespread—to a full blown characterisation of '#republic' as 'divided democracy', defined by polarisation, personalisation, social fragmentation of the public sphere through the echo chamber effect and 'cybercascades' of social and political influence.[17] Much of the scholarship directed at echo chambers is polemical (including Sunstein's argument). But where there is evidence based on available data, it suggests that algorithms play less of a role in exposure to attitude-challenging content than individuals' own choices and that 'on average more than 20 percent of an individual's Facebook friends who report an ideological affiliation are from the opposing party'.[18] The evidence further shows that those who do not use social media on average come across news from significantly fewer different online sources than those who do use social media[19] and that social media are actually reducing polarisation.[20] These findings suggest that although the design of platforms clearly affects the extent to which echo chambers exist, they are nowhere near as determined or 'hermetically sealed' as some critics suggest.

Similarly, for fake news, there is a lack of rigorous empirical evidence and it is also difficult to identify what should and should not be labelled as fake news, particularly given President Trump's indiscriminate and overuse of the term. Where there is data, it has suggested that on Twitter, false political news tends to spread more quickly and further than verified news (perhaps because of the novelty of false news) and that humans rather than robots were responsible for this faster spread.[21] Research has also started to suggest that the volume and effect of fake news has been exaggerated. Research in France and Germany found a very limited reach of identified disinformation providers on the open web, with only a few of them generating high levels of engagement on Facebook.[22] Evidence from the US 2016 presidential election, combining survey responses with individual-level web traffic histories, estimated that approximately one in four Americans visited a fake news website between 7 October and 14 November 2016. Trump supporters visited the most fake news websites, which were

overwhelmingly pro-Trump. In this way, there is a tendency for fake news to preach to the converted, reducing its overall effect; fake news consumption was heavily concentrated among a small group—almost six in ten visits to fake news websites came from the 10 per cent of people with the most conservative online information diets.[23] Likewise, another study examined online visitation data across mobile and desktop platforms in the months leading up to and following the 2016 presidential election and found that the fake news audience comprises a small, disloyal group of heavy internet users.[24]

For computational propaganda, there is little doubt of the scale, and it is being extensively gathered and analysed by the Computational Propaganda project at the University of Oxford.[25] An analysis of 9 million tweets from the 2016 presidential campaign found that one-third of pro-Trump Twitter traffic was driven by accounts that were bots or highly automated, compared with one-fifth of pro-Clinton traffic, and such an analysis has been performed for several elections since then. However, establishing the effect of such automated propaganda is far more difficult. This is especially the case with the now evidenced Russian activity in the 2016 election which, as a *Wired* article put it, was 'designed to look like it was coming from authentic American voices and interest groups', using automated accounts across platforms pushing out what was so-called native content—including video, visual, memetic, and text elements designed to push narrative themes, conspiracies and character attacks, rather than promoted or paid-for advertisements. This produced distorted social information, making it far more difficult for voters to assess where stories and narratives were coming from, whether they were real or propaganda, whether they represented the views of neighbours or not. Many were not actually aimed at changing opinion, but at reinforcing existing opinions and changing behaviour, making individuals feel strongly enough to do something. All were geared at disruption. As the *Wired* article noted, 'Persuasion and influence via social media cannot be estimated in linear terms; it requires looking at network effects. It is about the impact of a complex media environment with many layers, inputs, voices, amplifiers, and personalities. All of these elements change over time and interact with each other'.[26]

For hate speech, we know even less about the scale and scope of the phenomenon. Hate speech is clearly a problem that threatens to discourage a whole generation of women and ethnic minorities from public life, with misogyny towards campaigners for any issues connected with

women's rights particularly prominent in the UK. And social media platforms have been very slow to react, as highlighted by the tirades of abuse and hate directed at Caroline Criado Perez, the UK activist whose successful campaign in 2013 to have women represented on bank notes led to sustained and vitriolic harassment on Twitter, highlighting the problem for the first time and causing Twitter to change their complaints procedure. In 2016, several platforms (including Microsoft, YouTube, Facebook and Twitter) agreed a code of conduct with the EU to remove offensive material within twenty-four hours of reporting. According to EU figures, they were managing this in 81 per cent of cases by the end of 2017.[27] While there are qualitative studies of online misogyny[28] and new centres of quantitative research[29] focusing on hate speech, particularly on Twitter where data is available, we have little idea of the scale of hate speech, or the extent to which it relates to isolated individuals or is orchestrated by far-right groups, or the extent to which verbal threats relate to instances of actual hate crime (through 'doxxing', for example, where personal details and physical locations are revealed). Without such research, it is impossible to understand the extent to which the hate peddled by trolls is something new, somehow caused by the ease of expressing it through the availability of tiny acts of hate which then scale up, or whether it merely reveals society's dark secrets that were hitherto concealed as thoughts and never translated into any sort of action. Certainly, the apparent rise in hate is not restricted to its online form, as the campaigner Gina Miller, who instigated the Article 50 legal case against the UK government (and has received unprecedented levels of threats and harassment) put it:

> The idea that this abuse is the work of keyboard warriors is just not the case. These people take the time to make posters with vile images, put them in envelopes and post them. They go to the trouble of finding my email address or office number. This is really premeditated stuff.'[30]

We need the data to understand the claims of the anger-ridden and grief-stricken mourners of democracy in order to understand the scale and scope of hate and the mechanisms which drive it, and that means developing research partnerships with the most popular social media platforms.[31] Facebook

has now announced some positive moves, including the development of a data-sharing arrangement for reputable academic researchers, using a model developed by the Harvard political scientist Gary King and the Stanford lawyer Nate Persily.[32] They have also announced a new interface which will open up data on political advertising, the other threat to democracy posed by social media, in a move welcomed in May 2018 by the US *ProPublica* journalist Julia Angwin, who has done much to highlight the issue of highly targeted personalised political advertising. It is possible to discern a note of hope for the future role of social media in democracy from the ensuing exchange, which also highlights the importance of public service journalism in keeping up pressure, when on 11 May Facebook thanked Angwin (on Twitter) as follows:

Source: Twitter, 11 May 2018

Taking back control? Building transparency and accountability back into democracy

Any institutional catch-up for democratic systems to account for the new actors will need to be carried out at multiple levels; there is no simple solution or quick fix. First, as highlighted above, there is a role for public service journalism and social activism in keeping up the pressure on social media platforms to make more transparent their processes and algorithms which shape the information and news that people consume and use to make political decisions (such as whether to join a campaign and for whom to vote). From around 2013 onwards, social media platforms have taken a series of actions

to tackle the problems of hate speech, fake news and computational propaganda, and these moves have mostly been in response to public pressure. Facebook, for example, has employed thousands of fact checkers and moderators and in May 2018, published a transparency report[33] of the amount of content taken down on the site. The data highlights the scale of the challenge faced, with 3.5 million pieces of violent content and 2.5 million pieces of hate speech removed during the first quarter of the year, and the disabling of an astonishing 583 million fake accounts. The report illustrates the technical challenge and the importance of innovation in this area; while automated machine learning can deal with violent and even terrorist content relatively easily, hate speech is far more difficult to remove and only 38 per cent is dealt with automatically. And when it comes to the growing phenomenon of political bots directed by electoral campaigns or external forces wishing to disrupt elections, the answer may be a counter-spiral of automation by the internet corporations themselves. While an early example of this, the Microsoft 'Tay' chatbot driven by artificial intelligence, was a miserable failure (because it learnt quickly from the company it kept to spew out racist and abusive venom in all directions), more recent examples have been more successful.[34] Technologists and philosophers will have a role here, particularly in the building of ethical bots that do not go either native, or beyond their remit.

While there is a clear role for public pressure, regulation should not be removed from the policy table, and indeed when Mark Zuckerberg appeared before the European Parliament in May 2018, he made his first public acknowledgement that regulation was 'important and inevitable' and that Facebook was performing a public role. Although the German route of treating platforms legislatively like publishers (introduced in October 2017) is believed by many to be impossible to police,[35] the European Union's General Data Protection Regulation (GDPR) which came into force in May 2018 shows how regulation can effect meaningful behavioural change by commercial firms (as evidenced by the hundreds of emails and opportunities to discontinue relationships that everyone received in the preceding weeks). It also illustrated how the European state is leading in the way in the regulatory space.

There is a limit, however, to what regulation can achieve. In attempting to ban encrypted platforms such as WhatsApp in 2017, the then UK home secretary Amber Rudd declared that 'enough is enough'—an even more meaningless statement than 'Brexit means Brexit'. Such a move would

poison relations with, for example, Facebook while driving miscreants to far darker and harder-to-reach places, representing a massive act of environmental pollution. Another potential area for regulation is in challenging the monopolies of the huge platforms. The fact that Facebook was allowed to buy Instagram and WhatsApp, its key competitors, was a signal that the Federal Trade Commission (FTC) had totally underestimated the importance of this market. In the wake of the Cambridge Analytica affair, a new campaign, #freedomfromFacebook has emerged, for the purpose of lobbying the FTC to break up Facebook into these constituent parts. The political journalist Paul Mason has suggested that we should ferret around in the short history of social media to find those platforms, like Soundcloud, Medium and Twitter, that encourage co-operation and creativity and are 'worth saving', and establish new (co-operative) ownership models.[36]

We also need to stop denigrating tiny acts and extend our idea of what is democratic participation. For example, another way of tackling fake news may be to enlist the support of volunteers for fact checking and reporting false stories, in the same way that Wikipedia was formed. Confronting hate online at the individual level has been shown to be possible for some public figures, such as the TV academic Mary Beard, who confronted her trolls individually—and there may be creative ways of crowdsourcing such confrontations to avoid individual harm or risk. We also need to improve the education of children in terms of news consumption. It is by now clear that young people do not take news from conventional sources, as in national newspapers, TV or radio, where established brands are known as trusted sources, with flagged political standpoints and some kind of indicators of quality. News items on social media platforms, which is the only place where young people are likely to see them, could come from anywhere and are surrounded by advertising, and people navigate this environment with very little help.

Stabilising democracy in the social media age

Digital platforms now form the basis of our democratic environment and we must protect them rather than rushing to despairing of their influence as a terminal illness. While these platforms undoubtedly shape the information that we consume and the political decisions we make, the pathologies that they introduce are not terminal, but rather, chronic and under-researched, requiring careful study and long-

term management. We need to work out ways to take over the reins that at present, technology companies seem to hold—although their control is very much less absolute than we tend to believe. Any rethinking of democracy in the social media age must be multi-faceted, thoughtful, collaborative and evidence based. It will involve ethical and legal frameworks to guide as well as mandate good behaviour; working with tech companies rather than only making enemies of them; smarter policing of activities that are already illegal; and crowdsourcing safety in online spaces, so that people and social enterprises play a role.

This is not a sleek hyper-modernist vision of democracy, in spite of the high technology content. Indeed, by highlighting political turbulence and the messy chaotic nature of contemporary politics, I have tried to convey how an element of randomness has entered political life with social media. We have an understandable tendency to assume that because something important has happened, it was somehow inevitable. For example, the vote to leave the EU feels like the inevitable result of the clear pathologies of a divided country, where those with few resources had gained little from either Europeanisation or globalisation. Other seemingly landmark political events, such as the election of Donald Trump, have this same aura of inevitability, in spite of the surprise that greeted their arrival. Yet there is an alternative view. When we look in detail at the waves of support that led up to the vote for Leave, or Trump, or any of the other closely fought political contests with surprising results that have taken place over the past few years, perhaps it could have been different. The interconnectedness of our political life means that every tiny act of support for the Leave (and Remain) campaigns sent a tiny signal of viability out to other voters in a connected cluster of support—even if it was just a comment on a TV debate—which have scaled up to some kind of success, within some political micro-climate (such as a locality, or institution, or profession), but could have fizzled out into failure. Learning how to manage this unpredictability in democratic life is crucial for a stabilising of democracy in the social media age. This is a democracy built on workarounds and fixes, a messy solution for a disorganised, chaotic politics.

Notes

[1] S. Levitsky and D. Ziblatt, *How Democracies Die*, New York, Crown, 2018.

[2] 'The report of my death was an exaggeration', written by Mark Twain to Frank Marshal White, the English correspondent to the *New York Journal*, 31 May 1897, in response to rumours of his serious illness and death, and published in the journal on 2 June.

[3] 'Facebook now has 2 billion monthly users…and responsibility', Techcrunch.com, 27 June 2017; https://techcrunch.com/2017/06/27/facebook-2-billion-users/ (accessed 19 August 2018).

[4] 'Total number and the share of population of active social media and mobile social media users in the United Kingdom (UK) in January 2018', Statista.com;

https://www.statista.com/statistics/507405/uk-active-social-media-and-mobile-social-media-users/ (accessed 11 August 2018).

[5] M. Gladwell, 'Why the revolution will not be tweeted', *The New Yorker*, 4 October 2010.

[6] H. Margetts, P. John, S. Hale and T. Yasseri, *Political Turbulence: How Social Media Shape Collective Action*, Princeton, Princeton University Press, 2015. See this book also for analysis and visualisation of the data cited above.

[7] M. J. Salganik, P. S. Dodds and D. J. Watts, 'Experimental study of inequality and unpredictability in an artificial cultural market', *Science*, vol. 311, no. 5762, 2006, pp. 854–856.

[8] S. A. Hale, P. John, H. Margetts and T. Yasseri, 'How digital design shapes political participation: a natural experiment with social information', *PloS one*, vol. 13, no. 4, 2018, p.e0196068.

[9] For a discussion, see H. Margetts, 'Political behaviour and the acoustics of social media', *Nature Human Behaviour*, vol. 1, no. 4, 2017, s41562-017.

[10] Carole Cadwalladr made this point at an event hosted by Damian Collins MP and the Latvian Embassy on 'Resisting the onset of post-truth in democracy' in Parliament, 15 May 2018.

[11] E. Kübler-Ross, *On Death and Dying*, Abingdon, Routledge, 1969. Originally developed by Kübler-Ross to describe the emotions experienced by the dying, these five stages have been used widely to encapsulate the way in which people respond to bereavement and grief more generally.

[12] 'The Musgroves, like their houses, were in a state of alteration, perhaps of improvement', J. Austen, *Persuasion,* 1817, chap. 5.

[13] For a discussion, see D. Karpf, 'Online political mobilization from the advocacy group's perspective: looking beyond clicktivism', *Policy & Internet*, vol. 2, no. 4, 2010, pp. 7–41.

[14] J. Perry Barlow, 'A declaration of the independence of cyberspace', Electronic Frontier Foundation, 8 February 1996; https://www.eff.org/cyberspace-independence (accessed 11 August 2018).

[15] This is the view of *Spiked*, the advocacy group and magazine promoting free speech; http://www.spiked-online.com/ (accessed 11 August 2018).

[16] Definition of 'post-truth', designated word of the year in 2016, *Oxford English Dictionary*.

[17] The most recent of which is C. R. Sunstein, *#Republic: Divided Democracy in the Age of Social Media,* Princeton, Princeton University Press, 2017.

[18] E. Bakshy, S. Messing and L. A. Adamic, 'Exposure to ideologically diverse news and opinion on Facebook', *Science*, vol. 348, no. 6239, 2015, pp. 1130–1132.

[19] R. Nielson, A. Cornia and A. Kalogeropoulos, 'Challenges and opportunities for news media and journalism in an increasingly digital, mobile, and social media environment', Council of Europe Report, 2016; http://reutersinstitute.politics.ox.ac.uk/sites/default/files/research/files/Challenges%2520and%2520opportunities%2520for%2520news%2520media%2520and%2520journalism%2520in%2520an%2520increasingly%2520digital%252C%2520mobile%2520and%2520social%2520media%2520environment.pdf (accessed 11 August 2018); N. Newman, R. Fletcher, D. A. L. Levy and R. K. Nielsen, Reuters Institute Digital News Report, SSRN Scholarly Paper No. ID 2796534, 2016. M. Duggan and A. Smith, 'Political content on social media', Pew Research Social Science Research Network, 25 October 2016, also supports this view; http://www.pewinternet.org/2016/10/25/political-content-on-social-media/ (accessed 11 August 2018).

[20] P. Barberá, J. T. Jost, J. Nagler, J. A, Tucker and R. Bonneau, 'Tweeting from left to right: is online political communication more than an echo chamber?', *Psychological Science*, vol. 26, no. 10, 2015, pp. 1531–1542.

[21] S. Vosoughi, D. Roy and S. Aral, 'The spread of true and false news online', *Science*, vol. 359, no. 6380, 2018, pp. 1146–1151.

[22] R. Fletcher, A. Cornia, L. Graves and R. Nielsen, 'Measuring the reach of "fake news" and online disinformation in Europe', Reuters Institute factsheet, 2017; https://reutersinstitute.politics.ox.ac.uk/our-research/measuring-reach-fake-news-and-online-disinformation-europe (accessed 11 August 2018).

[23] A. Guess, B. Nyhan and J. Reifler, 'Selective exposure to disinformation: evidence from the consumption of fake news during the 2016 US presidential campaign', 2018; https://www.dartmouth.edu/~nyhan/fake-news-2016.pdf (accessed 11 August 2018).

[24] J. L. Nelson and H. Taneja, 'The small, disloyal fake news audience: the role of audience availability in fake news consumption', *New Media & Society,* February 2018, 1461444818758715.

[25] The Computational Propaganda Research Project, University of Oxford; http://comprop.oii.ox.ac.uk/ (accessed 11 August 2018).

[26] M. McKew, 'Did Russia affect the 2016 election? It is now undeniable', *Wired*, 16 February 2018; https://www.wired.com/story/did-russia-affect-the-2016-election-its-now-undeniable/ (accessed 11 August 2018).

[27] Reuters, 'Social media companies accelerate removals of online hate speech', 19 January 2018; https://www.reuters.com/article/us-eu-hatespeech/social-media-companies-accelerate-removals-of-online-hate-speech-eu-idUSKBN1F806X (accessed 11 August 2018).

[28] Qualitative studies include D. K. Citron, *Hate Crimes in Cyberspace*, Cambridge MA, Harvard University Press, 2014; K. Mantilla, *Gendertrolling: How Misogyny Went Viral*, Westport, Praeger, 2015. B. Poland, *Haters: Harassment, Abuse and Violence Online*, Lincoln NE, University of Nebraska Press, 2016.

[29] Such as the ESRC-funded Social Data Science Laboratory, at the University of Cardiff.

[30] L. O'Carroll, interview with Gina Miller, 'I've been told that "as a coloured woman", I'm not even human', *The Guardian*, 25 January 2017; https://www.theguardian.com/politics/2017/jan/25/parliament-alone-is-sovereign-gina-miller-speaks-out-after-article-50-victory (accessed 11 August 2018).

[31] One example is Denmark, where the government has appointed a 'digital ambassador', in an effort to build closer ties with internet giants such as Apple and Google.

[32] Facebook news, 'Facebook launches new initiative to help scholars assess social media's impact on elections', 9 April 2018; https://newsroom.fb.com/news/2018/04/new-elections-initiative/ (accessed 11 August 2018).

[33] Facebook news, 'Facebook publishes enforcement numbers for the first time', 15 May 2018; https://newsroom.fb.com/news/2018/05/enforcement-numbers/ (accessed 11 August 2018).

[34] V. Woollaston, 'Following the failure of Tay, Microsoft is back with new chatbot Zo', *Wired*, 6 December 2016; https://www.wired.co.uk/article/microsoft-zo-ai-chatbot-tay (accessed 11 August 2018).

[35] Free Speech Debate, 'The UK can show the way on platform regulation. But not by treating Facebook and Google as publishers', 6 November 2017; http://freespeechdebate.com/discuss/the-uk-can-show-the-way-on-platform-regulation-but-not-by-treating-facebook-and-google-as-publishers/ (accessed 11 August 2018).

[36] P. Mason, 'Why social media is like the railways—and must be saved', *The Guardian*, 9 January 2017; https://www.theguardian.com/commentisfree/2017/jan/09/social-media-fake-news-soundcloud-medium-facebook (accessed 11 August 2018).