DIGITAL THREATS TO DEMOCRACY

EXECUTIVE SUMMARY

The Workshop

The Law Foundation NEWZEALAND

Luminate

DIGITAL THREATS TO DEMOCRACY

INTRODUCTION

As we completed this report it was announced New Zealand's Prime Minister, Jacinda Ardern, would meet the French President, Emmanuel Macron in Paris to "bring together countries and tech companies in an attempt to stop social media being used to promote terrorism." The meeting will invite world leaders and tech company CEOs to sign a pledge called the 'Christchurch Call'.

The question is no longer whether something needs to change. The question has become: what precisely needs to change? And even more importantly: what can be done? What evidence do we have as to the interventions and solutions that might mitigate against the biggest threats posed to our democracy by digital media, without losing the best of the opportunities that the internet offers. Those are the questions we set about answering with this research.

One of the challenges of rapidly developing a policy response on digital media in response to a situation like the Christchurch attacks is that this entire area of policy has been relatively neglected until recently. As one participant in this research said, we need a better system for making policy on these issues before we can be any kind of global leader. In order to build our capacity as a country to understand and deal with these issues, we need a better evidence base.

What our research shows is that it is critical that the Prime Minister and her advisors look beyond immediate concerns about violent extremism and content moderation, to consider the wider context in which digital media is having a growing, and increasingly negative, impact on our democracy.

BACKGROUND

Over recent years a growing body of international research has looked at the impact of digital media on democracy, with particular focus on the US and the UK, where the role played by digital media in the election of Trump and the Brexit referendum raised significant concerns.

Our project was designed to find out if we should be worried about these same issues here in New Zealand, and if so, what should we do about it? In order to answer that question we identified five key features of democracy against which we could measure the impact of digital media, for better and for worse. They are:

> Electoral process and pluralism

> Active, informed citizens

> Shared democratic culture

> Civil liberties and competitive economy

> Trust in authority

WHAT WE'VE FOUND

Critically, we found that digital media is having an impact across every one of those features of a healthy democracy.

There are indicators that digital media has had some beneficial impacts. Our quantitative research here in New Zealand indicates, for example, that people from minority groups have been able to use digital media to participate in democratic processes including accessing political players, and engaging in public debate. Whatever our response to the challenges posed to democracy by digital media, it's important we don't lose these opportunities in the process.

But the overall trend should raise serious concerns. Active citizenship is being undermined in a variety of ways. Online abuse, harassment and hate - particularly of women, people of colour, queer people, people with disabilities and people from minority religions - undermines democratic participation not only online, but offline. Misinformation, disinformation and mal-information are undermining not only informed debate, but also public trust in all forms of information. Distraction and information overload are eroding citizens' capacity to focus on important and complex issues, and their capacity to make the 'important moral judgements' required in a healthy democracy.

Likewise, interviewees described a myriad of ways in which our shared democratic culture is being undermined by digital media - including through disinformation, polarisation, attention hijacking and radicalisation.

One of the clearest impacts of digital media on our democracy has been its impact on funding for mainstream media. While Facebook and Google hoover up the advertising revenue that once would have been spent on print, radio and television advertising, they contribute nothing to the work of producing the kind of news and current affairs reporting that is essential to a functioning democracy.

The representative survey we carried out indicates that New Zealand's small size and relatively healthy mainstream media (relative to elsewhere and despite significant resource

challenges) may help us avoid the worst effects of 'filter bubbles' and 'echo chambers' in digital media on some issues.

Interviewees in our qualitative research nonetheless pointed to examples where debate in New Zealand about issues like free speech, hate speech and gender identity attracted the attention of foreign actors holding strong, even extreme, views on these issues. Engagement by these foreign actors in the online public debates on these issues here in New Zealand appears to have contributed to a polarisation of views here.

THE THREE CORE PROBLEMS TO EMERGE FROM OUR RESEARCH

At the heart of the challenges to democracy posed by digital media are three core problems:

1. Platform monopolies: two or three corporations control not only our means of communication, but also the content which is distributed, both of which are core aspects of our democracy. Whilst the market power and global mobility of these companies make it possible for them to avoid national regulatory measures, either by moving operations elsewhere or simply ignoring them;

2. Algorithmic opacity: algorithmic engines are using huge quantities of personal data to make ever more precise predictions about what we want to see and hear, and having ever increasing influence over what we think and do, with little transparency about how they work or accountability for their impact; and

3. Attention economy: the dominant business model of digital media prioritises the amplification of whatever content is best at grabbing our attention, while avoiding responsibility for the impact that content has on our collective wellbeing and our democracy. The negative impact is brutally clear from both the literature and the world around us.

THE NEED FOR A SYSTEMIC RESPONSE

The key message is clear; digital media is having massive, system-wide impacts on our democracy. It affects every part of our lives and the people who run the corporations controlling the major platforms are having a determinative impact on the very structures and functions of our society. While better content moderation is clearly one of the responses we must demand of the platforms, it is not even close to being a sufficient response to the scale of the challenge.

It's critical that this moment of global cooperation is used to address the wider, structural drivers of the biggest threats posed to democracy by digital media. These structural drivers include the power that a handful of privately-owned platforms wield over so many aspects of our lives, from what information we see, who we interact with, and who can access information about us. And we must do this while maintaining and building upon the many opportunities digital media simultaneously offer, to tackle some of the biggest challenges facing democracy, including inequity of access and declining engagement.

THE OPPORTUNITIES OF DIGITAL MEDIA

The potential of digital democracy lies in its ability to increase democratic participation, embrace diversity of opinion, and empower marginalised groups. We identified six clear opportunities from the literature that digital media offers. These are: the democratisation of information publishing, broadening the public sphere, increasing equality of access to and participation within political processes, increasing participation and engagement in political processes, increasing transparency and accountability from government and promotion of democratic values.

Broadly speaking these opportunities fit into two categories: First, those that enable individuals, citizens, or groups, who due to their status in society have been excluded from fully participating in different aspects of the democratic process, through greater access to the levers of democracy. Examples include the use of digital media to: broaden the public's engagement with indigenous people and their lives, give more exposure to women in politics, build well-networked, educated and empowered communities, and encourage political engagement from youth.

The second category of opportunities relate to digital media's use by people in governments to make processes of democracy more inclusive, to increase engagement with citizens, improve transparency of government work, and rebuild trust in democratic processes. Examples of such work include: online deliberative democracy processes, open or e-government initiatives, and funding of public service journalism, platformed on digital media.

The opportunities for digital media are significant and important. If used well, digital media can enable governments to respond effectively to the experiences of marginalised groups, to ensure equitable policies and practices are designed, delivered and adjusted, and to build trust in the democratic institution as responding to the needs of all people. It offers as much to people pushing against barriers to their progression, inclusion, and improved wellbeing in society, as it does to people in government looking to remove those barriers and build a more inclusive democratic system.

THE THREATS

The threats to this promise outlined in the literature are significant however, and most are intricately bound up with the concentration of power in profit driven companies. The seven key threats we identified to inclusive democracy from digital media were: the increasing power of private platforms, foreign government interference in democratic processes, surveillance and data protection issues, fake news, misinformation and disinformation, filter bubbles and echo chambers, hate speech and trolling, and distrust/ dissatisfaction with democracy.

Researchers highlighted the increasing dominance of an increasingly small number of privately-owned platforms over the internet. People who own and control these platforms have a monopolisation tendency linked to the relationship between the mining of user-data and their imperative to make profits. This model of operation is termed "platform monopolies". The monopolisation tendency makes alternatives to the data-extraction for profit model, for example, co-operative, democratised ownership models, hard to start up and survive.

The concentrated power of these platforms shapes not just the wider information context and ability to develop alternative non-extractive models of digital information provision and sharing, but individual's personal lives also. Platform monopolies affects how we interact socially and with whom through algorithms. A body of literature points to how the actions of the people running these companies impact human rights, both through the control of personal data and the level of control over what appears in the public sphere.

From this model of platform monopolies flows a series of further threats to democracy. Some relate to the features of the platforms, directly linked to the capture of people's personal data. The collection and on-sale of personal data by these platforms, to both governments looking to undertake surveillance on their own citizens, and private organisations wishing to make profits, erodes trust in information systems by the public, and curtails the professional work of media and writers - a key plank in our democracy.

The creation of the "attention economy" also poses a significant threat. People's propensity to attend to shocking, false, or emotive information, especially political information, is exploited and used as a commodity product by digital media platforms. The literature shows that governments with the means and inclination to manipulate information can tailor false information towards individuals with the express intent of interfering in other countries democratic processes, for example the Russian government interference in the US election of 2016 using 'bots' and disinformation campaigns. While misinformation & disinformation, especially political disinformation, targeted at

individuals on digital media, is used to influence politics with a big and small p, from national elections through to information provision and sharing with regard to political issues and policy more generally.

Filter bubbles are a specific technical effect of this attention economy. Facebook's news feed is a filter bubble, created by a machine-learning algorithm which draws on data created by user networks, likes and comments, and how much organisations are willing to pay to be present there. Filter bubbles and the related echo chambers they feed into (in which people attend only to information which confirms what they already believe) are linked to a decline in trust in the ability to traditional news media to provide reliable information, have been found to exacerbate political divisions and polarisation, and has negative implications for the mechanisms of liberal democracy, as developing a broad consensus around decisions made in the public good becomes increasingly difficult.

The rise of hate speech and trolling is linked to the polarising effects of filter bubbles and echo chambers. Both racialised and sexualised hate speech is a specific threat exacerbated by the anonymity provided by digital media. Hate speech and trolling on digital media encourages affected groups to retreat to safe locations, rather than engaging with national debates and institutions. Research has found a correlation between strong, vocal disagreements with an individual's perspectives and a "spiral of silence" which acts to curtail the voicing of contentious opinions by minority groups. The particular ability of trolls and hate speech to fan antagonistic 'flames' rather than promote rational debates, has a direct impact on democratic participation.

While people's distrust with democractic process is a longer term issue, digital media has likely exacerbated this pattern across western democracies. Researchers argue that trust, informed dialogue, mutual consent, and participation - fundamental features of democracy - are being eroded by the features that make social media so profitable. Researchers also found that the way in which the information is distributed on digital media (horizontal, and decentralised and interactive) increases intolerance of others, polarisation and skepticism toward democracy.

The opportunities of digital media, while still apparent, appear to have been suppressed by the sheer weight of fake news, filter bubbles, populism, polarisation, hate speech, trolls, and bots, that have emerged from the concentration of power in a small group of private organisations seeking to maximise profits. Digital platforms initially celebrated for their democratic possibilities, have transformed into anti-democratic power centres through the collection and exploitation of users attention and data. These privately owned platforms have largely escaped public oversight or regulation over their ability to harness this new power for commercial or political gain.

The question is what can policymakers do to recalibrate? Are there empirically tested public policies and approaches that can ensure digital media works to strengthen and deepen democracy?

THE
SOLUTIONS

While our literature review was not exhaustive, in general we found a dearth of empirically tested solutions. Likewise, and possibly because of the lack of evidence in the literature, the experts interviewed for this research generally had more to say about the risks and threats they saw arising from digital media, than they did on potential workable solutions. However, we did find some common ground on solutions between the literature and in the interviews. Below we discuss potential solutions for each of the identified threats, with a focus on optimising the opportunities. We have organised those solutions in a hierarchy, from those we think will have the greatest impact, with the least effort required by individual citizens, through to those with the least impact and most individual citizen effort. That is not to say that the politics of implementing those solutions with the greatest impact are not difficult, but that the political effort required is justified by the potential for positive impact.

CHANGE SOCIETY WIDE STRUCTURAL & SYSTEMS ISSUES

This section focuses on structural and systemic change, addressing for example the disproportionate power of the tech giants vis-a-vis governments, citizens and their domestic competitors.

REDUCE THE POWER OF PRIVATE PLATFORMS BY:

Regulating platforms like other industries. Currently, regulatory debates largely centre around defining the structure, terms and conditions of what kind of industry private intermediaries represent. How platforms should be regulated or governed thus partly hinges on how these services are defined; for example, whether social media platforms are media companies, public spaces, utilities or some other service, largely informs how they can ultimately be governed. There is little or no empirical evidence to show how regulation in this area would or would not work, and therefore adaptive approaches to policy and regulation will be needed. This will involve ensuring that the impacts of

any change are regularly monitored and changes made as needed in response to those findings.

Introducing new modes of collective action. Under industrial capitalism we had collective bargaining, the strike, for example forms of collective action that were sanctioned by law and had the support of society that allowed people to tame capitalism with legal protection. In relation to digital media researchers suggest there are opportunities for more collective action both by tech workers, demanding for example more ethical design in the products they work on, and by digital media users.

COMBAT FAKE NEWS BY:

Supporting a vibrant and diverse media sphere. One that balances strong, independent and adequately resourced public service media with a non-concentrated commercial media sector. This is proposed but untested idea.

CREATE SUPPORTIVE ENVIRONMENTS & CONTEXTS - MAKING THE DEFAULT INCLUSIVE AND SAFE

REDUCE THE POWER OF PRIVATE PLATFORMS BY:

Designing new competitive digital media solutions. Disruptive technology is needed to forge an alternative digital future that in turn, facilitates a more democratic internet. This means the creation of platforms offering a different set of affordances (ie not those driven by platform monopolies).

REDUCE INTERFERENCE FROM FOREIGN GOVERNMENTS AND POWERS BY:

Designing new cybersecurity infrastructure and drawing on "big datasets" to review and assess electoral policies. The research in this area is also largely normative, and seems to generally prescribe such infrastructure and reviews will reduce threats to elections and other political processes.

ADDRESS SURVEILLANCE & IMPROVE DATA PROTECTION BY:

Regulating companies' information management practices. Some regulatory measures, like the Singaporean Data Protection Act 2012, work to and have been proven effective in bringing formal charges to data mismanagement and abuse.

Making regulatory changes to data privacy policies. However, there is little evidence to suggest that these changes will reduce surveillance/data collection so much as regulate how that data is stored, accessed and used by data collectors and other third parties

REDUCE THE POWER OF PRIVATE PLATFORMS BY:

Building citizen-consumer activism and creating a "sea change in public opinion".
Scholars and theorists suggest that a shift in public attitudes is needed to persuade digital media companies to change, there is however no empirical data to draw upon as to how effective this approach would or would not be.

COMBAT FAKE NEWS BY:

Developing and circulating persuasive counter-narratives. The focus would need to be on emotional not rational, appeal. This is proposed but unmeasured.

OVERCOME FILTER BUBBLES/ECHO CHAMBERS, AND SILENCING EFFECTS OF HATE SPEECH BY::

Supporting new platform designs with different design affordances.
The design of platform affordance has an impact on inclusion and participation, as well as the types of interactions people experience and information they are exposed to. There is some suggestion that design affordances can reduce the effects of filter bubbles by engaging internet users in more ideologically diverse communities. Well-designed, collectively-owned, online deliberative fora like Loomio have been empirically shown to also create a safe environment for marginalised groups. Research suggests that intentionally building more participatory forms of engagement into platforms might reduce filter bubbles, echo chambers and incivility, while increasing communication and deliberative processes.

IMPROVE TRUST IN DEMOCRACY BY:

The creation, selection and use of online platforms that afford citizen participation and deliberation. Some empirical evidence shows that direct and participatory democratic engagement/processes, e-government, and open government improve trust. International research has found that engaging citizens in deliberative processes often results in profound changes in deliberating citizens' "frequently in the direction of more common good-oriented policies."

Using digital government processes. Transparent, easy to access and well designed e-government and open government initiatives have been shown to increase positive feelings and citizen trust in local government.

CREATE LONG LASTING PROTECTIONS FOR PEOPLE OR INTERVENE TO PROTECT THEM

REDUCE THE POWER OF PRIVATE PLATFORMS, COMBAT INCIVILITY AND MISINFORMATION ONLINE BY:

Improving Content Moderation. Calls for new regulatory policies around content moderation acknowledge this remains an opaque and difficult practice, and on its own is not a fix-all solution. Current policies at the largest intermediaries attempt to balance stakeholder expectations (including users, consumers, advertisers, shareholders, the general public), commercial business goals, and jurisdictional norms and legal demands (which are generally governed by liberal-democratic (US) notions of "free speech"). Goals related to inclusive and participatory democracy are not included.

The most common 'workable solution' presented as it relates to content moderation, are processes that combine technical and social (human) responses. Advances in semi or fully automated systems, including deep learning, show increased promise in identifying inappropriate content and drastically reducing the number of messages human moderators then need to review. In the literature however, researchers note that neither automated nor manual classifications systems can ever be "neutral" or free from human bias. Human and/or automated content moderation is unlikely to achieve "civil discourse" or goals through moderation alone. Therefore, the combination of automated classification and deletion systems and human efforts remains the most effective content moderation strategy currently on offer. In the few places where they exist government regulations on private intermediaries' moderation practices have not been empirically tested for their efficacy or effectiveness.

COMBAT FAKE NEWS BY:

A multi-stakeholder content moderation. This is an approach that combines human and technical intervention, however this is a proposed but untested solution.

REDUCE HATE SPEECH/TROLLING BY:

Using identity verification systems. Sites that do not allow anonymisation and force pre-registration have been shown to solicit qualitatively better, but quantitatively fewer, user comments because of the extra effort required for engaging in discussion. Empirical research has also found that abusive comments are minimised when anonymous commenting is prohibited.

BUILD UNDERSTANDING AND CHANGE INDIVIDUAL BEHAVIOURS

ADDRESS SURVEILLANCE AND DATA PRIVACY ISSUES BY:

Encouraging individuals to employ technical solutions. Such solutions include ad-blockers and ad-tracking browser extensions, private browser options (e.g. Tor), open source platforms and cooperative platform models. "Evidence" supporting the efficacy of these tools and alternatives, however, is typically anecdotal.

COMBATE FAKE NEWS BY:

Education, particularly around critical thinking. Evidence has emerged in health for this approach.

REDUCE HATE SPEECH BY:

Building Resilience through Support Networks. Developing fast and effective reporting mechanisms and support networks. A networked approach can effectively combat the effects of hate speech; by building counter-narratives that counteract racism for example.

Coordinating diverse stakeholders to apply pressure to private intermediaries, in 'long-haul' campaigns, has also been effective in having hateful content removed from social media. Speed of removal is considered essential to diffusing the power of hate speech and trolling.

IMPROVE TRUST IN DEMOCRACY BY:

Civics education. Educating children in schools on 'good citizenship' has been positively associated with increased political engagement.

CONCLUSIONS

THE THREE CORE PROBLEMS TO EMERGE FROM OUR RESEARCH

At the heart of the challenges to democracy posed by digital media are three core problems:

> Platform monopolies: two or three corporations control not only our means of communication, but also the content which is distributed both of which are core aspects of our democracy, whilst the market power and global mobility of these companies make it possible for them to avoid national regulatory measures either by moving operations elsewhere or simply ignoring them;

> Algorithmic opacity: algorithmic engines are using huge quantities of personal data to make ever more precise predictions about what we want to see and hear, and having ever increasing influence over what we think and do, with little transparency about how they work or accountability for their impact; and

> Attention economy: the dominant business model of digital media prioritises the amplification of whatever content is best at grabbing our attention, while avoiding responsibility for the impact that content has on our collective wellbeing and our democracy. The negative impact is brutally clear from both the literature and the world around us.

KEY PRINCIPLES FOR POLICY RESPONSE:

Use democratic processes, which provide some degree of transparency about the decisions being made, accountability as to their impacts, and opportunities for challenge and judicial review. These processes must include meaningful participation by diverse representatives of the people whose lives are impacted by digital media. In particular, Internet users and civil society must have meaningful involvement, as the crucial third party in the multi-stakeholder process.

Draw on the evidence as to what is most likely to work, where it exists. Perhaps the most predictable finding of this research is that there has been little or no investment by people in government or other research funders into experimenting with and recording possible solutions, and there needs to be more.

Evidence-led and principled approach. Where there are gaps in the evidence, there are key principles that can be followed to reduce the risk of implementing solutions that do more harm than good. These include an evidence-led focus on 'upstream' structural change and the application of human rights principles.

Focus on structural or 'upstream' change. Tackle the structural drivers that underlie all the downstream problems - such as online abuse, disinformation, radicalisation and polarisation. Solutions should be designed to intervene at the structural level and to rebalance power through, for example: governance structures, regulation to restore transparency, accountability and fair competition and genuinely participatory and representative multi-stakeholder processes. None of this is to say that design solutions and platform affordances are not important. As the research shows, they will be essential. But without some rebalancing of power, without increasing the diversity of people involved in decision-making at the highest levels, those design solutions run the risk of replicating very similar problems to those we now face.

Respect and protect human rights. The following human rights principles should also be applied to policy development in this area:

> Universality: Human rights must be afforded to everyone, without exception.

> Indivisibility: Human rights are indivisible and interdependent.

> Participation: People have a right to participate in how decisions are made regarding protection of their rights.

> Accountability: Governments must create mechanisms of accountability for the enforcement of rights.

> Transparency: Transparency means governments must be open about all information and decision-making processes related to rights.

> Non-Discrimination: Human rights must be guaranteed without discrimination of any kind.

Agile approach. In the absence of a strong evidence base, it makes sense to take an agile, iterative approach to policy change. Experiment with all the policies all the time. Ensure that the funding, design, and implementation of policies reflect a record, learn, and adapt approach to measure the impact of any new initiatives or regulations, and to make adjustments as evidence becomes available as to impact.

URGENT AREAS FOR CHANGE

Some of the areas in which action is needed sooner rather than later include effort to:

Restore a genuinely multi-stakeholder approach to internet governance, including rebalancing power through meaningful mechanisms for collective engagement by citizens/users;

Refresh antitrust & competition regulation, taxation regimes and related enforcement mechanisms to align them across like-minded liberal democracies and restore competitive fairness, with a particular focus on public interest media;

Recommit to publicly funded democratic infrastructure including public interest media and the creation, selection and use of online platforms that afford citizen participation and deliberation;

Regulate for greater transparency and accountability from the platforms including algorithmic transparency and great accountability for verifying the sources of political advertising;

Revisit regulation of privacy and data protection to better protect indigenous rights to data sovereignty and redress the failures of a consent-based approach to data management; and

Recalibrate policies and protections to address not only individual rights and privacy but also to collective impact wellbeing. Policies designed to protect people online need to have indigenous thinking at their centre and should also ensure that all public agencies responsible for protecting democracy and human rights online reflect, in their leadership and approaches, the increasing diversity of our country.

ACKNOWLEDGEMENTS