Written evidence submitted by Reset (COR0141)

The nature and scale of the issue

1. The Covid-19 crisis has shone a spotlight on well-established problems in digital society. Rampant disinformation, fraud and profiteering, as well as the exploitation of vulnerable groups are just some of the ways in which modern technologies have exposed the public to added risk and harm in these times of public health emergency. These are not new problems, of course, but ones which have been compounded by our increasing dependence on digital media. And recent months have shown that the effects are not isolated or contained to one societal group. Rather, it is clear that these harms have widespread and dangerous - even deadly - implications for everyday life.

2. Worryingly, many legal and illegal online harms have proliferated in recent weeks, preying on people when they are most exposed. Vicious attempts to profiteer from the pandemic abound, with the Local Government Association noting how some councils have seen as much as a 40% increase in reported scams since the start of the crisis.[1] The elderly and vulnerable are particularly at risk[2], although the NCSC is being notified of hundreds of thousands of scam emails[3] by the broader public. Most notable are the callous attempts by criminals to exploit the coronavirus through fake offers of face masks and testing kits. The NSPCC has noted an increased risk of child online grooming during the lockdown[4] and the BMJ has raised concerns about gambling related harms and mental health during this unprecedented period.[5] Additionally, hate speech and extremist narratives are on the rise and well-documented by research from the ISD[6] and Centre for Countering Digital Hate.[7] The full spectrum of online harms is covered by Carnegie UK Trust's foundational work on harm reduction and the duty of care.[8] Their work is even more relevant in the context of Covid-19.

3. But perhaps the most prominent problem of recent months - which cuts across many other harms - has been disinformation. Ofcom's recent polling showed that exposure to disinformation is worryingly high, with 50% of respondents encountering false or misleading information on a weekly basis.[9] Notably, this extraordinary level of exposure is NOT declining week after week --

---

[1] Coronavirus: half a million shoddy face masks among surge in virus-related scams and illegal goods reported to councils, Local Government Association, 9 May 2020
[2] Rise in coronavirus fraudsters offering support to elderly for cash upfront, Local Government Association, 21 March 2020
[3] NCSC shines light on scams being foiled via pioneering new reporting service, National Cyber Security Centre, 7 May 2020
[4] Lonely children are twice as likely to be groomed online, NSPCC, 23 April 2020
[5] Covid-19: we must take urgent action to avoid an increase in problem gambling and gambling related harms, The BMJ, 6 April 2020
[6] Covid-19 Disinformation Briefing No.1, ISD Global, 27 March 2020
[7] https://www.counterhate.co.uk/
[8] Harm Reduction in Social Media, Carnegie UK Trust
[9] Covid-19 news and information: consumption, attitudes and behaviour, Ofcom, May 2020

despite much publicised attempts by tech companies to reduce disinformation. While it is encouraging that people feel able to spot fake news, much disinformation is, of course, hard to identify. The lines between fact and fiction are increasingly blurred. The result is, as the Ofcom polls show, that people are turning to traditional media outlets to access trustworthy information. However, over recent years, the business model of traditional public service journalism has been eroded and the professional newsrooms that we relied upon in the last century are fewer and farther between. The result is that the quality and quantity of mainstream news have both declined – particularly at the local level where trust in sources is often highest. There is less and less journalism to respond to public need in a crisis.

4. In place of traditional media outlets are the Big Tech firms, who sit at the centre of the debate about online harms and disinformation. Large technology companies (Google and Facebook in particular) now substantially control how information is organised and distributed in our societies. And their business model depends on maximising the amount of time users spend on their services (so that their attention can be sold to advertisers). As a result, the rise of digital platforms has given license to the loudest, coarsest, and most contentious voices in the public sphere to dominate social conversation online. This is because social media that optimise for capturing attention will tend to deliver the most sensational content. Falsehood, conspiracy, and outrage spreads faster than facts, complexity, and kindness. Covid-19 has only exacerbated this, as reports by the ISD[10] and Oxford Internet Institute[11] show. The countermeasures taken thus far by the industry cannot undo the fundamental logic of their businesses.

5. This relentless flow of information, and disinformation, is powered by machine-learning algorithms and automation. These algorithms serve the business model by maximising users' time spent on the platform - the number of clicks, shares, and pageviews that drive advertising sales. In a kind of pincer movement on the user, the social media platforms curate noncommercial (i.e. organic) information flows to optimise for engagement even as they sell targeted advertising designed to maximise influence over preferences. The unimaginable size of the training data, the world-leading investment in AI experts, and the scale of the testing interface delivers efficient results and ever-increasing quarterly returns for the Silicon Valley giants. As we spend more time online during the Covid-19 lockdown[12], these algorithms are being fed with even more data.

6. To their credit, tech firms have made unprecedented efforts to counter the false and misleading information cascading over their systems in recent weeks. Whatsapp's implementation of forwarding limits[13]; Facebook and YouTube's deletion of conspiracy theorist material[14]; and Twitter's labelling of misleading information[15] are just some of the ways in which

---

[10] Covid-19 Disinformation Briefing No.1, ISD Global, 27 March 2020
[11] Social Media Misinformation about the WHO, Au/Bright/Howard, 20 April 2020
[12] Locked-down households using internet for 41 hours a week, USwitch, 5 May 2020
[13] About forwarding limits, WhatsApp
[14] Facebook and YouTube are rushing to delete "Plandemic," a conspiracy-laden video, Technology Review, 7 May 2020
[15] An update on our continuity strategy during COVID-19, Twitter, 16 March 2020

Big Tech have tried to tackle the rise of disinformation. But it isn't enough. Disinformation circulated so rapidly during the early weeks of the pandemic, that firms are struggling to catch up. Content promoting conspiracy theories harmful to public health is being viewed millions of times before it can be removed or flagged as untrustworthy, only to reappear elsewhere in the digital realm.[16] NewsGuard's recent data highlights how certain Facebook Pages are being used as "super-spreaders" of Covid-19 disinformation, amplifying fake news across the platform: 36 European Facebook Pages alone reach 13,223,446 Facebook users.[17] Similarly, NewsGuard identified 10 Twitter accounts which are "super-spreaders" of disinformation, reaching over 3 million followers between them. Two of these fountains of fake news have even been validated by Twitter as verified accounts.[18] Social media is awash in links to clickbait sites that are profiting through online ads from people panic-consuming online information as they search for answers. Often these ads are served up by the same social media giants responsible for enabling the spread of harmful content.[19] And in many cases, the damage can't be undone - once a false narrative is unleashed, it is very difficult to unravel.

7. Indeed, the Covid-19 crisis has shed new light on the pervasiveness of Big Tech. In a time of social distancing, social media fills the void of human interaction more than ever. These services are fundamental in maintaining an engaged, informed and safe society, and, at present, we cannot co-exist without them.

The need for regulation

8. Well-intended attempts to counter disinformation should not absolve Big Tech from their broader responsibilities, or reduce government's expectations of them during or after the crisis. If anything, the past few months have demonstrated that these firms and services should be subject to even more public oversight, given the fundamental role they play in a functioning democracy and society. Watering-down or delaying upcoming legislation, such as the Online Harms Bill, will mean that these harms will continue to proliferate unabated (despite the efforts of the companies at marginal self-regulation). Government must develop and enact a public policy agenda that regulates the digital marketplace to align its interests with those of democratic and social integrity. Other complex, concentrated industries such as pharmaceuticals, manufacturing and financial services are overseen by government regulations. Digital platform giants, with all of the social impact we can see and that which we cannot, should not be omitted from the list. And, as recent surveys by doteveryone[20] and 5Rights Foundation[21] show, citizens want and expect tech firms to be regulated.

9. The UK's admirable efforts in this space should not lose momentum. The current crisis only strengthens the need for swift regulation - there is hardly a better case in point than public

---

[16] How the 'Plandemic' Movie and Its Falsehoods Spread Widely Online, New York Times, 20 May 2020
[17] Facebook 'Super-spreaders': Europe, NewsGuard, 5 May 2020
[18] Twitter Super-Spreaders, NewsGuard, 5 May 2020
[19] AdTech + Brands on Coronavirus Disinfo Sites: 11-15 May, Global Disinformation Index, 15 May 2020
[20] People, Power and Technology: The 2020 Digital Attitudes Report, doteveryone, May 2020
[21] 5Rights Foundation poll conducted by YouGov, June 2019

health disinformation and conspiracy to warrant these kinds of intervention. EU regulators have responded as such, demanding detailed information from Big Tech about how they are tackling disinformation during the pandemic.[22] Other regulators should feel empowered to follow suit.

Regulatory access to data, audit and inspection

10. The public policy agenda must be implemented in ways that preserve freedom of expression, privacy and security -- even as they protect the public from the myriad harms described above. At the center of the policy solution, we must focus on the mechanism that drives the digital media economy -- data. Companies should be required to share certain data about the design of their platforms with regulators, governments and academia. Without access to the data and AI systems that guide information flows in these markets, there is no obvious way to make good policy that will be adaptive and durable as the industry evolves. None of the issues at the centre of this debate can be adequately addressed without this auditing function: democratic election integrity, child online safety, anti-competitive practices, consumer fraud and abuse, harassment and hate speech, and much more. There need to be new systems for transparency and auditing of algorithmic design and decision making, giving regulators the powers and tools to inspect these powerful lines of code.

11. The presence of harmful content in the public sphere is nothing new. What is new is the application of AI to the business of information distribution and targeting that enables forms of artificial amplification of harmful content, increases the chance of high frequency exposure to extreme views, and opens up media channels that double as interpersonal communications networks to organised exploitation. It is these processes that must be examined and regulated to comply with standards consistent with democratic values.

12. The Online Harms White Paper identified this problem, stating that regulators should be able to “require additional information, including about the impact of the algorithms” and to “request explanations about the way algorithms operate”. This does not go far enough. Regulators need to have the tools and powers to test the operation of algorithms and to undertake inspections themselves. At present, there is a massive asymmetry of information. The harms are easily observed as specific incidents, and they do in fact appear to form a pattern. But the companies that hold the data that could verify these patterns and measure their scope hold all the data, and they do not make it available for independent review under any circumstances. This lid is kept tightly shut. Without access, regulators are forced to rely on the companies to police themselves through ineffective codes of conduct. This is extraordinary. We have an industry operating in markets with clear externalities that cause public harms. The companies have all the data and tools needed to track, measure and evaluate these harms - indeed these tools are a core part of their business. But they make none of these available to public oversight, even as they avoid all but the most basic interventions to protect the public from harm.

---

[22] EU demands tech giants hand over data on virus disinformation, Financial Times, 18 May 2020

13. There is precedent in the UK for a regulator to have such powers of oversight. The Information Commissioner's Office (ICO) has licence to undertake consensual audits to assess how data controllers or processors are complying with good practice in the processing of personal data.[23] Should the company not agree to a consensual audit, the ICO can seek a warrant to enter, search, inspect, examine and operate any equipment in order to determine whether a company is complying with the Data Protection Act.[24] Similarly, the Investigatory Powers Commissioner's Office (IPCO) has powers[25] to conduct investigations, inspections and audits as the Commissioner considers appropriate for the purpose of the Commissioner's functions, including access to apparatus, systems or other facilities or services.[26] \

14. Any future regulator of online harms will need a similar ability to carry out an algorithm inspection with the consent of the company; or if the company doesn't provide consent, and there are reasonable grounds to suspect they are failing to comply with requirements, to use compulsory audit powers. The resource to carry out these investigations could sit within the regulator, but they could also have the power to instruct independent experts to undertake an audit on their behalf. This would help ensure that the correct expertise is acquired for the work as is needed. This would mirror the Financial Conduct Authority's power to require reports from third parties; what they dub "skilled persons reviews".[27]

15. In addition, as recommended by the Centre for Data Ethics and Innovation, academics should be able to access certain datasets when conducting research into issues of public interest.[28] Efforts in this area are underway[29], but they have been challenging to establish, are limited in scope and are yet to prove themselves. While the online harm regulator will be able to "encourage" companies to give researchers access to data, its powers will need to go beyond mere encouragement. The power of these datasets should, in certain circumstances, be available to serve the wider public good. A joint paper prepared by Demos, doteveryone, Global Partners Digital, Institute for Strategic Dialogue and Open Rights Group, coordinated by Digital Action, provides more detailed recommendations.[30]

The need for a broad response

16. As the government understands, relying on the companies to remove or moderate harmful content is not the solution to the problems we face. The status quo doesn't work, as recent efforts have shown. There needs to be a much more systemic approach, proactively redefining how trustworthy information is accessed and distributed. To that end, the Online Harms Bill will

---

[23] s129, Part 5, Data Protection Act 2018
[24] Schedule 15, Data Protection Act 2018
[25] s235(1), Chapter 1, Part 8, Investigatory Powers Act 2016
[26] s235 (4), Chapter 1, Part 8, Investigatory Powers Act 2016
[27] Skilled person reviews, Financial Conduct Authority
[28] Review of online targeting, Centre for Data Ethics, February 2020
[29] https://socialscience.one/
[30] Algorithmic Inspection and Regulatory Access, Demos, doteveryone, Global Partners Digital, ISD, Open Rights Group and Digital Action, May 2020

be most effective if supported by a range of other interventions and initiatives. For example, the ICO should be sufficiently resourced and empowered to hold the technology sector to account.[31] More broadly, at a market level, governments need to modernise competition policy to reduce the power of monopolies. At a societal level, governments need to invest in education and digital news literacy, reestablishing some norms about how to discern credible and non-credible information online. And at a democratic level, governments need to make commitments to public service journalism and community media that provide alternatives to the mindless stream of click-bait and propaganda that dominates digital media today.

About Reset

Reset (www.reset.tech) was launched in March 2020 by Luminate in partnership with the Sandler Foundation. Reset seeks to improve the way in which digital information markets are governed, regulated and ultimately how they serve the public. We will do this through new public policy across a variety of areas – including data privacy, competition, elections, content moderation, security, taxation and education.

To achieve our mission, we make contracts and grants to accelerate activity in countries where specific opportunities for change arise. We hope to develop and support a network of partners that will inform the public and advocate for policy change. We are already working with a wide variety of organizations in government, philanthropy, civil society, industry and academia.

May 2020

[31] Fears the ICO is too feeble to take on social media giants prompts audit, The Telegraph, 9 May 2020