Full Version

Confronting Bad Information

Lies, half-truths, false narratives and unfounded conspiracy theories are today rampant on the Web and social media. It is difficult to put a figure on the net cost to society — the damage to the reputation of those targeted and their ability to function, the financial losses incurred, the risk to democracy, social cohesion, public safety, privacy and people's mental health, and the opportunity cost associated with trying to counter the threat. One study came up with a figure of $78 billion a year.[1]

The good news is that thousands of individuals and many hundreds of organisations and agencies around the world are today working on the problem. But the picture is complex and confusing and it is difficult to keep abreast of the rapid changes taking place, both in the technology (including the new threats that are emerging microtargeting, deepfakes, GPS Spoofing, etc.), and the politics / legislation. There's no shortage of information, indeed, quite the reverse: the infosphere is awash with reports on 'fake news' and conspiracy theories, and hundreds of books and reports have been published. So, for those who are struggling to make sense of it all, here's an overview of the runners and riders. For convenience, I've presented these under six headings and commented on some noteworthy initiatives.

I should also point out that some organisations are working on multiple fronts, and that there is a great deal of interaction / sharing of information and expertise between and within the different groups, including subcontracting and commissioning work and providing grants.

The purpose of this paper is not to attack social media or the Internet, which are integral parts of today's digital economy, rather it is to explore what efforts are being made to tackle threat from bad information / wild conspiracies and minimise the harm, and also to comment on the challenges that lie ahead — I've relegated an analysis of 'The Problem' to an Annex.

1 Fighting Fake

1.1 Individuals & NGOs

Many individuals and Non-Governmental Organisations are working to raise public awareness of 'fake news' (*DeSmogUK* & *StopFake*), denialism (*Debunking Denialism*) or online safety (*Internet Watch Foundation*), or monitoring domains for the risk of spreading bad information[2] (*Global Disinformation Index*) or the overall health of the Internet (*Mozilla Foundation*).[3] Other NGOs are tackling cyber bullying (*Civilination*) or hate speech (*#JaGarHar*), or fighting to protect people's privacy or freedom of speech (*Big Brother Watch, Electronic Frontier Foundation* & *Open Rights Group*); or promoting sound science (*Bad Science Watch*) or quality information (*Ethical Journalism Network* & *First Draft*). And we should not forget the countless cartoonists and satirists who have their own distinctive way of holding liars, sociopaths and charlatans to account (see e.g. Ziglis & Cooper).

There are also in excess of 300 fact-checking organisations working to "increase the cost of lying," including >60 that are verified signatories of the *International Fact Checking Network*. Last year's 'practitioner's conference', '*Global Fact',* was held online for the first time, significantly extending its reach. The same happened with *MisInfoCon,* a grassroots global movement focused on "building solutions to online trust, verification, fact checking, and reader experience in addressing misinformation in all of its forms." Then there's *Wikipedia* (the 13th most popular website), which has many pages devoted to mis / disinformation, conspiracy theories etc. Twice a year *Wikipedia* calls on librarians and "anyone with a passion for free knowledge" to support articles with missing references.[4] Countless webinars are also taking place, spreading the word, and volunteer experts on misinformation and deception are on hand to give advice (see e.g. the *Atlantic Council*'s *Portal*,[5] *Forum on Information & Democracy* and *Global Experts on Debunking of Misinformation*).

1.2 Schools & Colleges:

Children need to be able to distinguish between genuine and fake / fact and opinion / evidence and argument. Finland sets a good example: after being confronted by increased Russian trolling, President Niinisto reformed the education system to prioritise digital education, media literacy and critical thinking, and he called on every Finn to take personal responsibility for the fight.[6] And teachers can take comfort for the fact that there's now a great deal of educational material online.

---

1 There's a discussion of the challenge of costing the damage from bad information on the *Fighting Fake* website.

2 I'm defining 'bad information' as a combination of *misinformation* — information that is false, but not created with the intention of causing harm; *disinformation* — information that is false and deliberately created to cause harm; and *mal-information* — information intended to be private/confidential that is made public with the intention of inflicting harm (including 'doxxing').

3 Mozilla has just released its 4th Internet Health Report (for 2020).

4 It is germane to note here that *Wikipedia* talks, not about 'truth', but 'verifiability'.

5 The *Atlantic Council* also supports the *Digital Forensics Research Lab,* runs regular webinars, and holds an annual *Open Source Summit* [360/OS] to bring together its network of 'Digital Sherlocks' (journalists, activists & strategists) to share ideas and experiences..

6 Today Finland excels in league tables on happiness, press freedom, transparency and social justice, and gender equality (1st, 2nd, 3rd & 4th respectively).

1.3 Academics, Consultants & the Media

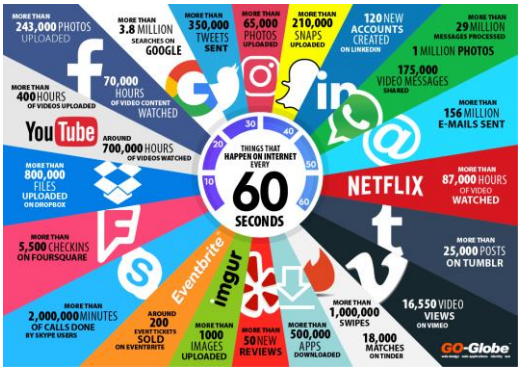
A growing number of universities are actively involved in researching into bad information. For example, in the UK there are teams in Cambridge (*Centre for Research in Arts, Social Sciences & Humanities*), Cardiff (*HateLab*), Cranfield (*National Cyber Deception Laboratory*), Goldsmiths (*Centre for Investigative Journalism*), the LSE (*Media Policy Project*) and Oxford (*Oxford Internet Institute* & *Reuters Institute for the Study of Journalism*). Academics are also researching conspiracy theory at Bath, Bristol, Cambridge, Essex, Kings College and Warwick (and doubtless elsewhere).

Many mainstream media organisations have a proud history of promoting factual reporting / quality journalism, and with the rise of 'fake news', some like *The New York Times* and *The Guardian* have been rewarded by increased subscriptions. The *BBC* (which has an *Anti-Disinformation Unit* & *Reality Check* Team) is involved in *Project Origin* (which seeks to "create a process where the provenance and technical integrity of content can be confirmed, establishing a chain of trust from the publisher to the consumer"). It has also linked up with tech firms to create an early warning system which alerts others to specific disinformation, and has set up a joint online media education campaign, which is co-operating on voter education and shared learning, especially around elections. Similar linkups can be found in the *European Union* (*Fighting Disinformation*), USA (*NewsQ*), and the Philippines (*The Consortium on Democracy & Disinformation*).[7]

Consultancies are also playing their part, for example, by championing the accurate analysis of data (*Gapminder*); harnessing digital forensics / looking for aberrant or suspicious behaviour on social media (*Bellingcat* & *Graphika*); tracking deepfakes (*DeepTrace*), or monitoring the impact of bad information on human rights (*Global Partners Digital*). Other groups have developed Internet tools such as *Botometer* (to identify bogus *Twitter* pages); *Fakespot.com* (spot fake reviews); *Junkipedia* (enable journalists/researchers to collect, track, analyse bad information spreading online); *Tineye* (reverse image searching), and there are many more.

1.4 Tech Giants

The *Tech Giants* were late to publicly acknowledge the problem of false accounts and fake, extremist or illegal material on their platforms, but in recent years they have been forced to respond to serious criticism in the press, confront some awkward realities (including claims of PTSD amongst content moderators (see 2.2 below) and even complicity in genocide) — and pay substantial fines for breaking local laws.[8] Today *Facebook* has a 'War Room' (to respond in real time to disinformation campaigns) and a number of 'Deletion Centres'; and it has been increasing its use of artificial intelligence to remove harmful or illegal content, and in 2019 it set up an independent Oversight Board to adjudicate on complaints.[9]



You get an indication of the size of the problems faced by social media from this graphic, which shows the level of traffic on the most popular platforms. In *2016* *Facebook* was dealing with 243,000 photos uploaded and 70,000 hours of video watched *every minute*. And since then the number of internet users has grown by a *third.*

1.5 National Governments

Politicians and regulators also have a responsibility for seeing that online disinformation doesn't cause harm to the public / organisations and undermine democracy. In the UK a host of agencies regulate service providers and NGOs and uphold the law, standards of public decency and intellectual property rights. These include the *Information Commissioner's Office* (ICO), *Ofcom*,[10] the *Health & Safety Executive*, *Charity Commission, Advertising Standards Authority* and *Crown Prosecution Service* (CPS).[11] Information about the work of the intelligence and security services (*GCHQ, MI6, MI5* & *National Cyber Security Centre*) is understandably sketchy, but *GCHQ* is reported to have shut down much of *ISIS's* online

---

7 Autocrats have different rules. For example, Russia's approach to information has been characterised by one *European Commissioner* in the following terms: "the truth is what people believe" and "there are no facts only interpretations." Comment by Véra Journová, *Commissioner for Justice, Consumers & Gender Equality* during an *Atlantic Council* Fireside Chat [8Dec2020]

8 The EU has been most proactive with its GDPR, legislation which stipulates fines of up to €20 million for violating its rules, or 4% of the annual worldwide turnover of the preceding financial year, whichever is greater. A recent example is the *Commission Nationale de l'Informatique et des Libertés* fining *Google* and *Amazon* a total of €135 million (around $163 m) for flouting France's data protection laws.

9 This initiative this falls someway short of what some critics are demanding who have recently formed what they call the '*Real Facebook Oversight Board*' which has a rather wider brief, content moderation, policies and a range of other platform issues. The unprecedented action taken by *Twitter* to ban President Trump from its platform (and *Google* and *Amazon* to close down *Parler*), may be signalling recognition that regulation of social media is overdue and likely to be a priority for the Biden administration.

10 It is reported that under the Online Safety [Online Harms] legislation currently going through Parliament *Ofcom* will be empowered to enforce a new legal 'duty of care' and force platforms like *Facebook* to remove 'harmful' content (e.g. terrorism, child abuse, self-harm and suicide imagery etc.). It is assumed that this will include bad information.

11 The ICO was set up to "uphold information rights in the public interest, promote openness by public bodies and data privacy for individuals." Anyone posting wrong or malicious information about someone online can be reported to the Commissioner, who may in turn take up the case with the CPS.

propaganda, and more recently, been targeting Covid-19 anti-vaccine initiatives linked to Russia. In respect of taking *offensive* action to combat disinformation / election interference, *GCHQ* has a *Joint Threat Research Intelligence Group* and is partnered with the *Ministry of Defence* in the new *National Cyber Force*. And the *British Army* has been upgrading its *Psychological Warfare Unit* and reconfiguring its *6th Division* to fight cyber threats.

A variety of parliamentary committees provide oversight, for example, the *Digital, Culture, Media & Sports Select Committee* has been particularly active and has produced several excellent briefings/reports, including 'Disinformation and Fake News' (which *inter alia* argues for social media platforms to be held responsible for harmful content on their services, and technology companies to be taxed to fund a public information campaign). Committees in the *House of Lords* have also produced impressive reports, for example on social media, the Internet and artificial intelligence.

1.6 Organisations and Agencies

A range of international organisations and agencies, both formal and informal, keep a watchful eye on Internet development and governance. These include the *Internet Society*, *Internet Governance Forum* and *IoT Global Council*;[12] other units are concerned with security/monitoring threats and identifying and exposing mischief-makers and bad information. In this respect, it is worth mentioning:

- *EU vs Disinfo*, a specialist unit set up in 2015 by the *European Union* to debunk stories put out by *The Kremlin* or pro-Kremlin proxies — to date >10,500 fake stories (details on the website); and
- the 'Verified' initiative, launched by the UN, which aims to "flood digital space with facts amid the Covid-19 crisis".

2 The Challenge Ahead

The development of the Internet raises a host of difficult and as yet unresolved issues.

2.1 Internet Conundrums

Here are six tricky Internet conundrums:

- Should the Internet be regulated to reduce misuse, disinformation and criminal activity, and if so how and by whom?
- Should big tech be broken up to increase competition and encourage innovation, if so, how?
- How should society weigh the benefits of anonymity and encryption against the social, economic and political costs?
- Should netizens have rights, including the right to own their data and have protection from surveillance capitalism / prying eyes? If so, how might they be policed?
- Should access to the Internet be a basic human right, and if so, how might this be achieved? (Access to the internet is *already* a human right in half a dozen countries.)
- What can be done to maintain the Internet as a global resource and prevent its balkanization into a 'splinternet'?

2.2 Other Challenges

And here are some other challenges:

For politicians and big tech:

- tackling the diverse range of threats from bad information *without* degrading internet services and compromising people's privacy, security and freedom of expression;
- regulating online content and policing conflicting narratives — *who* exactly should adjudicate what appears on platforms, or *what* is true / acceptable? It can't be (as of now) monopolies operating without democratic oversight.[13] And how should humour and biting satire be handled?[14]
- reducing the considerable human cost of viewing and moderating disturbing content on social media.[15]

Some regulatory measures that have been tried have had unintended and unwelcome consequences, not least driving people onto 'alt-tech' platforms, such as *Gab*, *Telegram* and *8kun*[16] and putting a chilling

---

[12] We are increasingly dependent on the *Internet of Things* (IoT) and *Bodies*, and this has led to widely expressed concern about potential security flaws, which makes people vulnerable to hacking and manipulation.

[13] Were *Facebook* and *Twitter* within their right to suspend Donald Trump from their platform following the violent white supremacist riots that breached the US Capitol? Yes, but it's a bit late! As Jillian York argues: "Trump has the First Amendment right to spew deranged nonsense, [but] so too do tech companies have the First Amendment right to remove that content." Both platforms appear to have stuck to the tenet that "content posted by elected officials deserves more protection than material from ordinary individuals, thus giving politicians' speech more power than that of the people" (a position York notes is at odds with evidence that hateful speech from public figures has a greater impact than similar speech from ordinary users). "The problem is that these policies have not been applied evenly around the world." Iran's Supreme Leaders and India's Narendra Modi (and many other demagogues) have not had their *Twitter* accounts shut down...

[14] Poe's Law is relevant here: "No matter how bizarre, outrageous, or just plain idiotic a parody of a Fundamentalist may seem, there will always be someone who cannot tell that it is a parody, having seen similar REAL ideas from real religious/political Fundamentalists."

[15] In May 2020, *Facebook* agreed to pay $52m to content moderators in the US as compensation for PTSD and other mental health issues they developed — some 11,250 moderators are eligible.

[16] Yasmin Green [Director of Research & Development at *Google's Jigsaw*] notes that 'alt-tech' platforms "market themselves as being

effect on freedom of expression. The latter happened with Germany's landmark 2017 legislation on hate speech,[17] which led to social media platforms becoming more cautious about what they would allow to be uploaded. Autocratic governments have also copied the legislation.

Important to inject here the notion that 'freedom of speech' does *not* mean 'freedom of reach' — or, for that matter, 'freedom to hurt' / 'freedom of consequences'.[18] And it is the platform's invisible algorithms that control the reach... As Renee DiResta points out "There is no right to algorithmic amplification. In fact, that's the very problem that needs fixing."

For *agencies, NGOs and academics*, there is the challenge of:

- reconciling the different goals, procedures and organisational cultures in order to facilitate more effective cooperation.[19]

And for teachers and concerned citizens, the challenge of:

- keeping abreast of information, often of dubious quality posted on websites or spread on social media and working out what's true anymore; and
- trying to "make sense of complexity in a world where human emotions, diverging interests and conflicting values coexist and intermingle with evidence"[20] whilst navigating the 'news' minefield which is laced with bad information and driven all too often by deception and bad intent.

The problems are multi-dimensional and difficult for most people to comprehend, especially given the speed at which some technologies (like AI) are evolving. The purpose of the *Fighting Fake* website is to try to provide a '*One Stop Shop'* where activists, educators and concerned citizens can find an overview of the issues and a discussion of what can and is being done to tackle them. [21]

And here are a few other observations:

- It's an old truism that technology can be used for good or for evil. But what makes regulation / policing the Internet so difficult is that the bad actors invariably seem to be ahead of the game. *Facebook, Google, Microsoft, Amazon* and *Reuters* have recently joined forces to fight one particularly disturbing development, the malicious use of deepfakes (discussed in the Annex), but "any defence can be countered by the next software update."
- The attack on the US *Capitol* on 6th January has been described as the "9/11 moment of social media". It has highlighted the growing pressure for Section 230 of the 1996 US *Communications Decency Act* to be reformed. This is the legislation that protects internet platforms that carry third-party content, the so-called 'safe harbour' rule. But as one commentator has argued: "If Facebook wants to boot Trump—or photos of breastfeeding mothers—that's the company's prerogative. The problem is not that Facebook has the right to do so, but that—owing to its acquisitions and unhindered growth—its users have virtually nowhere else to go and are stuck dealing with increasingly problematic rules and automated content moderation.[22]
- Die-hard conspiracy theorists are difficult to reach (or convince): some circulate fake or misleading materials in order to promote a deeper 'emotional truth', and without regard for the 'facts'. For many, *QAnon*, *AntiVax* of '*Stop the Steal*' are their proud 'badge of identify'; for others (like Alex Jones of *InfoWars*) conspiracy theories provide an opportunity to make money. Lots of it. How far technical measures will ever be able to counter unsupported allegations or blatant opportunism is not clear.

'anti-censorship' (read: unmoderated) and for 'free speech' (read: hate speech welcome)." Moving dangerous content in this way presents two dangers: first, "that the inaccessibility of these subjects could encourage people to seek them out. Spreaders of conspiracy theories will have the opportunity to sensationalise a 'censored' video ('what they don't want you to see') — what is known in online marketing as a 'curiosity gap';" second, and ironically, "the displacement of [such] ideas to more fringe platforms could help them spread, removing scepticism they may face from mainstream audiences." Jan Buts makes a similar point: "To combat the spread of anti-vaccination rumours, platforms are currently using a dual strategy of censorship and fact checking. Both practices have their pitfalls. Censorship may actually stimulate curiosity, while people who distrust mainstream media are not likely to trust fact checkers."

17 The *Netzwerkdurchsetzungsgesetz* (*NetzDG*) law was passed in June 2017 and came into force in early October. Complaints must be checked immediately, 'obviously illegal' content deleted within 24 hours, and complainants informed immediately of decisions taken. Fines of up to €5 million may be imposed if rules are infringed.

18 For the record: "The First Amendment to the United States Constitution guarantees free speech, and the degree to which incitement is protected speech is determined by the imminent lawless action test introduced by the 1969 Supreme Court decision in the case Brandenburg v. Ohio. The court ruled that incitement of events in the indefinite future was protected, but encouragement of 'imminent' illegal acts was *not* protected."

19 Many fine words are spoken and written about partnership, but there's not always a willingness to collaborate / share, perhaps because of the desire for publicity / acknowledgement, or simply the competition for resources...

20 Quote by Anthony Gooch [Director, *OECD Forum*] from a paper on 'Fighting Disinformation: A key pillar of the COVID-19 recovery'.

21 *Fighting Fake* maintains a database of organisations and initiatives for tackling bad information, and this has now reached well over 700 entries — and this is only in the English-speaking world. We are also working to identify the bad actors, both the amateur and the professional disinformation spreaders.

22 The observer, Jillian York, is Director for International Freedom of Expression at the *Electronic Frontier Foundation*. She argues that "Platforms needn't be neutral, but they must play fair". The "answer is not repealing Section 230... but in creating the conditions for more competition... It's not difficult to see how ratcheting up platform liability could cause even more vital speech to be removed by corporations whose sole interest is not in 'connecting the world' but in profiting from it." Note: In Oct 2020 a major congressional anti-trust report (backed by Democratic lawmakers) called for changes that could lead to the break-up of some of America's biggest tech companies. It followed a 16-month congressional investigation into *Google, Amazon, Facebook* and *Apple*.

3 The Battle for Truth / Democracy

Much has been written about how best to tackle bad information / wild conspiracies. Here's a compilation of proposals:[23]

3.1 Establish authoritative oversight bodies

Establish authoritative oversight bodies to identify generic problems with bad information, develop strategic solutions and foster cooperation between nations and organisations. We need bodies that can coordinate policy, challenge legislators' timidity, and help formulate new systems of regulation supported by sanctions that really hurt.

3.2 Dismantle the digital monopolies & introduce competition

Dismantle the digital monopolies and introduce competition and tackle social media's aggressive business models — this means reining in surveillance capitalism and microtargeting,[24] and making the platforms more transparent and accountable. Measures proposed include: formulating a new category where big tech combine the functions of platform and publisher (which could involve platforms taking responsibility for *amplifying content* even if they are not held responsible for the content itself); developing a global code of ethics which sets down in writing what is and what is not acceptable on social media, with significant financial penalties for companies that transgress; and ensuring paid-for political advertising data on social media platforms is transparent.[25]

3.3 Prepare Citizens for the Digital World

Alert citizens of all ages to the opportunities and dangers of the digital world and prepare them for what is to come (i.e. through public education and encouraging media literacy / critical thinking — see next item).

3.4 Refit Democracy for the Digital Age

To be *really* effective, there needs to be an *integrated approach* to the threat across all organisations / platforms, both within nations[26] and internationally. Governments must ensure the same rights online as offline, with strong data protection laws for personal data and clear rules on who's responsible when data is moved from one service to another. And social media platforms should have a legal duty to inform users of their privacy rights, especially with regard to profiling and automated decision-making.[27]

3.5 Improve Mainstream Media and Access to Trusted Information

The mainstream media needs to be safeguarded and supported and access to trustworthy content improved (with bad information, if not removed completely, made less visible and forced down the newsfeed). [28]

3.6 Take Cyber-Security and Information Warfare Seriously

in our digital world, cyber-crime is on the increase,[29] and information is being weaponised on an

---

23 Most of the ideas in this list are not original; they have been proposed by: Sandra Matz, Guy Rolnik & Moran Cerf, Carl Miller, Renee DiResta & Mike Godwin or taken from reports cited elsewhere in the piece.

24 Microtargeting is used by political parties (and candidates) to track individual voters and identify potential supporters — and sometimes to discourage those thought likely to vote for the opposition.

25 This needs to explain clearly what data is collected and how it will be used. Big Tech companies should: state their terms and conditions — and in language children can understand; embrace 'freedom of thought' as a policy commitment and perform due diligence on how their activities may harm it; and release regular transparency reports which explain how they are tackling hate speech and mis/disinformation.

26 Finland has already been mentioned. One of its near neighbours, Lithuania, has also had issues with Russian disinformation, which led to it setting up *Demaskuok* ('Debunk'), a counter-disinformation campaign supported by more than 4,000 'elves' — volunteer journalists, IT professionals, businesspeople, students and scientists. The elves routinely scan articles in the media/on the Internet against a database of trigger words and narratives and their findings are sent to interested parties (NGOs, newsrooms, politicians, etc.) and the *Defence Ministry* produces regular written or video 'debunks' for the public. Taiwan's approach is also interesting as it demonstrates the country's ability to harness the power of its civil society and tech industry through a robust public-private partnership initiative. Its success in getting disinformation under control is also due to "government crackdowns on groups that spread disinformation, Taiwan's initiatives to improve media literacy, and President Tsai Ing-wen's decision to prioritize the problem, exemplified by her appointment of Audrey Tang, a software engineer, as digital minister in 2016."

27 When it comes to transparency, Norway has a novel approach: by law every person's tax return is browsable by the public, with just one caveat. If you view somebody else's tax return, they'll receive a message telling them. In other words: information is in the public domain, but so are your actions, should you choose to behave inappropriately.

28 The *Cairncross Review* (Feb 2019) proposed the establishment of an *Institute for Public Interest News* dedicated to amplifying "efforts to ensure the sustainability of public-interest news." It called *inter alia* for online platforms to set out codes of conduct for commercial agreements with news publishers (with agreements approved and overseen by a regulator "with powers to insist on compliance"); a market study of the online advertising industry; new forms of tax relief on digital publications; support for public interest journalism; and an expansion of the local democracy reporting service (ultimately managed by the new Institute).

29 Cyber-crime accounts for more than 50% of reported crime in UK (this is probably an underestimate as much crime goes unreported). An FBI report found that (in 2019 and excluding the US), the UK was top of a list of 'Top 20 Victim Countries, with 93,796 casualties, far more than the second on the list, Canada, with 3,721.

industrial scale by domestic and foreign actors. It has proved to be a highly effective/damaging and relatively cheap to produce; moreover, bad information is easy to circulate without getting caught (if you know what you're doing). Politicians clearly need to be more proactive in responding when interference is detected and countermeasures implemented in order to make attacks more expensive, both politically and economically.[30] Security treaties also need to be updated to recognise this new form of attack.[31]

As Joseph Nye[32] points out: "In the information age it's not just whose army wins but whose story wins." Russia currently tells the best stories, but China is fast catching up, and not far behind them is a melee of domestic and international extremist groups and organised criminals.

4 Conclusion

> "While once social media was seen as a liberating means to speak truth to power, now the issue is how to speak truth to social media." Wael Ghonim[33]

So, we should take comfort from the fact that the severity of the threat posed by bad information / wild conspiracy theories has been recognised and that a formidable network of individuals and organisations are now working on The Problem. We cannot expect a quick fix or 'silver bullet'. We should also take seriously widespread concerns about the lack of coordination between the major players, which is a consequence of diverse interests and conflicting political agendas and legislative approaches — not to mention the very real prospect of Internet balkanisation and the capacity of bad actors for 'malign creativity'. There's also the conundrum of how best to regulate big tech without compromising or curtailing the many services on which we are now so reliant. All these issues are rising up the international agenda — witness the way discussion of disinformation / the infodemic has crept onto the agenda of bodies such as the *Global Economic Forum*, *NATO*, *OECD*, *UNESCO* and *World Health Organisation*.

But perhaps the most important message to take away from this discussion is that bad information is *everybody's* problem and that we all need to play our part by holding to account those who spread half-truths, lies and misinformation. We can best do this on social media and at the ballot box. We can also subscribe to quality publications and support *bona fide* groups that are fighting fake and those politicians brave enough to take a stand. Few things can surely be more important for social stability and peaceful co-existence than having a common understanding of facts and events, and what is 'true', and speaking out when these critical pillars of our democracy are misrepresented or ignored.

Dr Mike Flood *Fighting Fake* [28th January 2021]

About

As an activist and campaigner, I have been concerned about bad information since before 'fake news' raised its ugly head in the runup to the 2016 US Presidential Elections. I set up *Fighting Fake* just after President Trump was sworn into office with the aim of helping raise public awareness of the threat. For those interested there's more information about this/me on the *Fighting Fake* website.

Annex: Bad Information

Bad information is generated by a broad range of bad actors and spread via the internet and social media, often using encryption and the Dark Web (to hide identity/cover tracks).

A1 Bad Actors

At the one end of what AC Grayling calls "the biggest lavatory wall in history"[34] there are the thoughtless or ill-informed individuals who pass on clickbait and tittle tattle without a thought for the consequences; then there are the entrenched conspiracy theorists, many motivated by bizarre convictions, and those out for revenge; and the petty criminals and fraudsters. And, at the other end of the 'wall', there are the organised criminal gangs, dishonest or deceitful politicians/cheerleaders, single-minded fanatics, and hostile foreign powers. Their combined efforts have helped:

- undermine public trust in science, government, the media, business and civil society;

---

30 This could involve suspending biased media channels (like RT); publicising illicit activities; freezing oligarch's assets / restricting their travel [eg 2010 US *Magnitsky Act*]; and or sanctioning perpetrator's goods and services.

31 Failure to act may be because international law, or actions / attribution is ambiguous, or the impact does not justify a response (as discussed in the 2013 'Tallinn Manual'). Perhaps worth noting here that the US military is reported to have blocked Internet access to the infamous Russian 'troll factory' (the *Internet Research Agency*) in St Petersburg (reportedly run by one of Putin's close associates) to prevent it sowing discord among Americans during the 2018 mid-term election.

32 Nye coined the term 'soft power' — getting others to want the outcomes that you want. The Russians call this 'reflexive control', and they are very good at it, with their coordinated black propaganda and disinformation.

33 Wael Ghonim, who helped ignite the Egyptian Arab Spring with his Facebook campaign.

34 Other unflattering descriptions of the Internet include: "an electronic asylum filled with babbling loonies" (Mike Royko) and "democracy's revenge on democracy" (Molly Haskell).

- damage economic prospects, confidence and morale;
- destabilize the political process and undermine democratic government; and
- put lives at risk, not least by increasing the political tension between different camps / nation states and compromising our ability to tackle existential global threats, not least pandemics and the climate crisis.

Bad information is not a new phenomenon. However, in recent years several factors have combined to make it a more potent force in the world including the development of the Internet, mobile technology, social media, computer power / big data and AI. We have seen:

- a massive growth in 'fake news', hate speech, cybercrime and disinformation;
- social media hijacked by populist politicians, crooks and foreign powers to manipulate opinion and sow division;
- personal information harvested on an industrial scale and exploited or sold on; and
- failure to regulate 'surveillance capitalism' and the big tech companies that thrive on it.

Indeed, 'fake news', mischief-making, deception and lying have become an integral part of our 'post truth' world where the line between fact, opinion and belief has become blurred, facts are used selectively, sometimes mischievously, evidence and reasoned analysis are ignored, and expert opinion dismissed or reviled;[35] and fabricated stories and staged events are designed to play to specific audiences and or mislead the public. I've summarised these issues below[36] and explained why the threat is not going to go away anytime soon.

A2 Anatomy of the Problem

The problem with bad information is multi-dimensional. It encompasses:

- *Broken Public Trust* — bad information undermines public trust, confidence, morale and can lead to serious harm.
- *Genuine or Fake?* — knowing whether information is genuine or fake can be highly problematic, especially when aspects of a story are true.
- *Sticky Memes* — fake information is 'stickier' than real news; it can be produced anonymously and at little cost; and it spreads significantly faster.
- *Anti-Social Media* — social media platforms promote and amplify ill-informed or malicious voices and inflate 'likes'/the audience.
- *Tech Giants* — tech giants have from the start adopted a strategy of anything goes until it is pronounced illegal. They have shown themselves unable or unwilling to purge their platforms of bogus web sites / fake, extremist or illegal material, which today thrives on their platforms despite their attempts to stop it.
- *Conspiracy Theorists* cause confusion, anxiety and polarisation by their toxic mixture of misleading, manipulated and/or fabricated content and fake science.
- *Information Warfare* — extremists and hostile states take advantage of liberal democracies to manipulate private data / use bad information to damage markets and social cohesion, and discredit democracy.
- Lack of Regulation & Coordination — regulating the tech giants / online content is proving highly problematic, not helped by internet breakup and poor coordination amongst groups fighting fake.

Each of these problems needs to be addressed in a different way, by different actors and over different timescales.

Worth noting here that government-backed disinformation campaigns are on the rise: a new report from the *Oxford Internet Institute* finds that online campaigns designed to discredit opponents, influence public opinion, drown out dissent and meddle in foreign affairs were waged in no less than 81 countries in 2020, up from 28 three years ago. The authors describe this as an 'industrial scale problem'.[37]

A3 New & Developing Threats

The development of the Internet raises profound and difficult questions which society is still struggling with:

---

[35] It is perhaps worth pointing out here that one of the *positive* things to have come out of the Covid-19 experience is that the public today better understands our dependence of science and expert advice. (In the UK, faith in 'experts' was seriously undermined by politicians' outpourings during the rancorous Brexit debate, most notably Michael Gove's comment that "people in this country have had enough of experts")

[36] There's a more in-depth discussion of the issues (and the routine abuse of personal data by big tech) on the *Fighting Fake* website.

[37] The *OII* report focuses on the use of 'cyber troops' (teams from the government, the military or political parties which are committed to manipulating public opinion on social media) which regularly conduct 'computational propaganda' campaigns involving the "use of programmed bots or humans to spread purposefully misleading information across the internet." It also found "an alarming increase in the use of 'disinformation-for-hire' services across the world. Using government and political party funding, private-sector cyber troops are increasingly being hired to spread manipulated messages online, or to drown out other voices on social media."

- the routine collection and manipulation of our information via smart devices in our homes, offices and cars;
- the covert use of Facial Recognition Technology for surveillance and other purposes, and the prospects of mind-reading;
- the synthesis / manipulation of voices, photos, video and text including 'deepfakes';
- the manipulation of satellite imagery [including *Google Maps*] and GPS Spoofing; and
- China's expected dominance in AI, and its growing political confidence and global influence.

With deepfakes we don't know who produces, but we know whose agenda they serve; and because they are so realistic, deepfakes can scramble our understanding of truth. And as we become more attuned to their existence, our trust in the veracity of *all videos* will be undermined, including those that are genuine.

How far the threat posed by deepfakes and the other developments will get worse before it can be reined in is an open question. This depends on so many considerations, not least the development of detection technology and effective countermeasures, carefully formulated regulation measures, and the extent to which the public understands and takes heed of the threat. The new Cold War between the US and China doesn't help.