

BIG BROTHER WATCH

**Submission to the Science,
Innovation and Technology
Committee's Inquiry: Social
media, misinformation and
harmful algorithms**

December 2024

About Big Brother Watch

Big Brother Watch is a civil liberties and privacy campaigning organisation, fighting for a free future. We're determined to reclaim our privacy and defend freedoms at this time of enormous technological change.

We're a fiercely independent, non-partisan and non-profit group who work to roll back the surveillance state and protect rights in parliament, the media or the courts if we have to.

We publish unique investigations and pursue powerful public campaigns. We work relentlessly to inform, amplify and empower the public voice so we can collectively reclaim our privacy, defend our civil liberties and protect freedoms for the future.

Contact

Silkie Carlo

Director

Direct line: 020 8075 8478

Email: silkie.carlo@bigbrotherwatch.org.uk

Madeleine Stone

Senior Advocacy Officer

Direct line: 07864733785

Email: madeleine.stone@bigbrotherwatch.org.uk

Introduction

We welcome the opportunity to provide evidence to the Science, Innovation and Technology Committee's important inquiry into social media, misinformation and harmful algorithms.

Big Brother Watch has campaigned extensively on online safety and speech issues since the Online Harms White Paper was published in 2019. We have produced several reports on freedom of expression online, as well as briefing policymakers and parliamentarians as the Online Safety Act progressed through Parliament.¹ We believe the Government's role should be focused on upholding the rule of law online while protecting its citizens' rights to freedom of expression and privacy. We have long-standing concerns that attempts to legislate to restrict 'misinformation' online would likely have grave consequences for free speech online and would be open to political exploitation.

In response to this call for evidence, we will respond thematically, addressing definitions of misinformation and the problems this poses for freedom of expression, social media business models and algorithms, the current legal framework for combatting misinformation online and the work of the National Security Online Information Team ('NSOIT').

Defining misinformation and problems for freedom of expression

As an organisation that campaigns for freedom of speech online, we are concerned by the ongoing implications of attempts to regulate 'misinformation' and online 'harm' for free speech. There can be no doubt that there is a considerable amount of content online that is untrue and unpleasant. However, where such content does not breach criminal law, government and social media platforms' interventions to 'moderate' speech can interfere with the right to freedom of expression a right which is protected in UK law by the Human Rights Act 1998 and European Convention on Human Rights.

One of the primary concerns with attempts to combat online misinformation and disinformation arise from the difficulty in defining these terms. While disinformation is defined by the UK government as "deliberate creation and spreading of false and/or manipulated information that is intended to deceive and mislead people, either for the purposes of causing harm, or for political, personal or financial gain", and misinformation is generally seen as the inadvertent spread of false information, these terms are expansive, and highly subjective.² The UN special rapporteur on freedom of

¹ See Ministry of Truth: The secretive government units spying on your speech (2023) and The State of Online Free Speech (2021): <https://bigbrotherwatch.org.uk/reports/>; See briefings on 'Online regulations' and 'Free speech online': <https://bigbrotherwatch.org.uk/research/>

² Preventing misinformation and disinformation in online filter bubbles – House of Commons Library, 15 January 2024: <https://commonslibrary.parliament.uk/research-briefings/cdp-2024-0003/>

expression noted “the concept is undefined and open to abuse” and that “the lack of consensus [on how to define mis/disinformation] underlines the complex, intrinsically political and contested nature of the concept.”³ Amnesty International has noted that countries across the world have introduced laws to clamp down on ‘fake news’ and ‘misinformation’, including Cambodia, China, the Gulf Countries, Indonesia, Nicaragua, the Philippines, Russia, with each country defining the terms differently. They found that these laws have been:

“a tool for controlling what is discussed in public and gives the authorities the power to censor uncomfortable information and determine what is considered to be true or false, offensive, dangerous or seditious in a way that enables them to target dissenting and critical voices”⁴

The significant range in how these terms are defined is evident in Ofcom’s recent research, *Understanding misinformation: an exploration of UK adults’ behaviour and attitudes*:

*“It is important to be clear from the outset that misinformation is a subjective term.
(...)”*

Perceptions of misinformation include and aren’t limited to: provision of empirically false information; provision of information that someone doesn’t agree with; provision of information that doesn’t fit with someone’s prior knowledge of, or existing beliefs about, a subject – which can result in true information being reported as false, and vice versa; and something that a public figure has said and is being reported on by a news platform or service – with such reporting either identifying the statement as misleading, or providing it as if it were accurate.”⁵

In the political context, all major parties have been accused of spreading misinformation or disinformation by political opponents and campaigners, typically in relation to complex policy areas which are the subject of considerable debate.⁶ Clearly, individuals and organisation take very different positions on what constitutes

3 Disinformation and freedom of opinion and expression: Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression – Human Rights Council, Forty-seventh session, A/HRC/47/25, 13 April 2021: <https://documents.un.org/doc/undoc/gen/g21/085/64/pdf/g2108564.pdf>

4 A Human Rights Approach to Tackle Disinformation: Submission to the Office of the High Commissioner for Human Rights – Amnesty International, 14 April 2022: <https://www.amnesty.org/en/wp-content/uploads/2022/04/IOR4054862022ENGLISH.pdf>

5 Understanding misinformation: an exploration of UK adults’ behaviour and attitudes – Making Sense Of Media, Ofcom, 27 November 2024: <https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/media-literacy-research/making-sense-of-media/dis-and-mis-information-research/mis-and-disinformation-report.pdf?v=386069>

6 See for example: CCHQ, X, 3 April 2024: <https://x.com/CCHQPress/status/1775472980442714210>; Reform UK, X, 14 November 2024: https://x.com/reformparty_uk/status/1857095023478645197; Tories accused of false claims on new 20mph limit by minister – BBC News, 12 September 2023: <https://www.bbc.co.uk/news/uk-wales-politics-66739588>; Swinney spreading ‘misinformation and lies’ about Labour’s plans, warns Sarwar – PA Media, 13 June 2024: <https://www.msn.com/en-gb/news/uknews/swinney-spreading-misinformation-and-lies-about-labour-s-plans-warns-sarwar/ar-BB1oaxne>; The Labour Party, X, 20 January 2012: <https://x.com/UKLabour/status/160301819409608705>; General election 2019: Ads are ‘indecent, dishonest and untruthful’ – BBC News, 17 December 2019: <https://www.bbc.co.uk/news/technology-50726500>

misinformation, and often define it in opposition to their existing beliefs or political position. It is not possible, or desirable in a democratic society, for any institution, organisation or government to act as an absolute arbiter of truth.

It is also important to stress that, under human rights legislation, expression should not be restricted based on its veracity alone. In *Salov v. Ukraine*, the European Court of Human Rights found:

*"Article 10 of the [European] Convention [on Human Rights, on freedom of expression] does not prohibit discussion or dissemination of information received even if it is strongly suspected that this information might not be truthful."*⁷

Where citizens do not incite violence, defame others or infringe upon other rights, the right to speak freely has never been a right conditional on its relative truth. One of the many reasons that freedom of speech is vital to the continuation of a free society is that it allows ideas to be developed through discussion and debate. Censorship on the basis of accuracy alone threatens to stymie the epistemic process, which allows knowledge to develop and sees it put to the test in a democratic and free society. Professor Frank Kelly, author of the Royal Society's report into scientific misinformation, made this point when cautioning against the censorship of scientific content on the basis of its accuracy:

"Science stands on the edge of error and the nature of the scientific endeavour at the frontiers means there is always uncertainty.

(...)

*"This is important to bear in mind when we are looking to limit scientific misinformation's harms to society. Clamping down on claims outside the consensus may seem desirable, but it can hamper the scientific process and force genuinely malicious content underground."*⁸

During the Covid-19 pandemic, concerns around scientific misinformation were heightened and led to censorship decisions which ultimately contradicted developing public health advice. For example, in March 2020 Facebook announced that since January 2020 it had "banned ads and commerce listings that imply a product guarantees a cure or prevents people from contracting COVID-19."⁹ Whilst there was little transparency about precisely how Facebook judged content falling into this category, the company publicised the fact that under the policy, adverts for face masks

⁷ *Salov v. Ukraine*, ECtHR, application No. 65518/01, judgment, 6 September 2005, para 113

⁸ Royal Society cautions against censorship of scientific misinformation online – The Royal Society, 19 January 2022, accessed 2 December 2024: <https://royalsociety.org/news/2022/01/scientific-misinformation-report/>

⁹ Combating COVID-19 Misinformation Across Our Apps – Nick Clegg, Meta, 25 March 2020: <https://about.fb.com/news/2020/03/combating-covid-19-misinformation/>

were banned.¹⁰ This rush to remove ‘misinformation’ contradicted later public health messaging.

As former Supreme Court judge Lord Sumption wrote, critiquing a particularly censorious version of the government’s Online Safety Bill:

“We have to accept the implications of human curiosity. Some of what people say will be wrong. Some of it may even be harmful. But we cannot discover truth without accommodating error. It is the price that we pay for allowing knowledge and understanding to develop and human civilisation to progress.”¹¹

Social media business models and algorithms

It is widely recognised that the online data trade means many social media companies’ terms and conditions are primarily designed for their own economic benefit and legal protection rather than to protect the interests of their users. In what has been termed ‘surveillance capitalism’, social media companies profit by extracting vast quantities of highly sensitive personal data from their users, and utilising this data to micro-target content and advertising back to the same users.¹² The business model of social media companies is built from mass scale violations of privacy and data protection law across the globe.¹³ Not only does this produce vast profits for Big Tech companies, it also gives them an extraordinary level of control over the flow of information and ideas in the internet age.

Online platforms are designed with the intention of keeping users on the site for as long as possible. This allows platforms’ systems to gather more data on users while keeping users exposed to digital adverts based on information that the site garners about them. This feedback loop itself is protected by amplifying content that catches the user’s eye. Algorithmic content ranking promotes posts based on users’ preferences which can perpetuate personal biases but also gives more weight to content which is controversial or incendiary. Indeed, Ofcom noted that “virality and algorithmic recommendations” played a role in the spread of inflammatory content related to the Southport stabbings earlier this year.¹⁴ This is an issue concerning design and not one which can or should be fixed by inhibiting free speech online.

10 Banning Ads – Combating COVID-19 Misinformation Across Our Apps – Meta, 19 march 2020: <https://about.fb.com/news/2020/12/coronavirus/#banning-ads>; Allowing the Promotion of Non-Medical Masks on Facebook – Meta, 10 June 2020: <https://www.facebook.com/business/news/allowing-the-promotion-of-non-medical-masks-on-facebook/>

11 The hidden harms in the Online Safety Bill – Jonathan Sumption, The Spectator, 20 August 2022: <https://www.spectator.co.uk/article/the-hidden-harms-in-the-online-safety-bill>

12 The term was coined by Shoshana Zuboff in her influential book The Age of Surveillance Capitalism, 2019

13 Complaints under the GDPR against the online advertising industry have been filed in the UK, Ireland, Poland, Spain, the Netherlands, Belgium and Luxembourg, concerning breaches of data protection law: <https://fixad.tech/ad-tech-gdpr-complaint-is-extended-to-five-more-european-regulators/>

14 Ofcom: Clear link between online posts and violent disorder – BBC News, 24 October 2024: <https://www.bbc.co.uk/news/articles/c70w0ne4zexo>

The solutions often offered to combating misinformation are de facto censorship, be that de-ranking certain content, 'shadow banning' users or removing posts. This approach grants social media platforms enormous power over the public's freedom of expression. Rather, a focus on the design and processes that platforms use to keep users scrolling is a more fruitful and rights-respecting approach. Algorithmic transparency is crucial, in order to allow policymakers and civic society the opportunity to examine and scrutinise how these vast corporations operate. This kind of transparency should also extend to content moderation and platforms should be willing to publish granular data both on any actions taken on users' posts and on the broader functionality of their algorithms.

Platforms should consider how content is ranked in a way which does not promote sensationalism. Platforms should also make algorithmic systems open to public scrutiny.

Further, companies should work to further expand and simplify user controls over the information they see, the people they are exposed to, and the recommendations they are shown. Unlike the physical world, users can exercise considerable control of the information and views they are exposed to online by blocking others, muting key words, controlling news feeds, and using age-appropriate controls. User control helps people to mitigate the subjective "harm" they might otherwise be exposed to. This approach protects freedom of expression in our online public squares whilst allowing people to create diverse experiences that reflect their own preferences, interests and needs.

Lastly, we welcome initiatives to promote digital literacy – although we believe this is a role for our national education system rather than for tech companies. Digital literacy, combined with more effective user controls, would allow individuals to take better control of their online experiences.

Current legislative framework

The Online Safety Act ('OSA') passed into law on 26 October 2023, and seeks to regulate the online sphere by placing new duties on social media companies and search services. Throughout the passage of the Act, we stressed that provisions in the Act will have significant ramifications for the online speech of UK citizens and would empower Big Tech companies to act as privatised speech police. We remain concerned about the impact the Act will have on the right to privacy and freedom of expression.

The independent regulator for Online Safety, Ofcom is tasked with implementing the Act's provisions by producing Codes of Practice and guidance on how service

providers can comply with their online safety obligations and assessing the extent to which platforms fulfil their regulatory obligations. The proposed timeline for the “illegal harms” duties to come into force is March 2025.¹⁵ The child protection safety duties are due to become enforceable in July 2025.

Given that the Act’s provisions have not yet come into force, calls for further restrictions on social media platforms and online speech related to misinformation and disinformation are both premature and will frustrate Ofcom’s ongoing consultation work in relation to its implementation of the existing legislation.

Provisions on disinformation

Section 179 of the OSA provides a new false information offence. An individual commits this offence if they send a message – without reasonable excuse – which they know to be false and which they intended to cause non-trivial psychological or physical harm to a likely audience. This offence replaced section 127(2) of the Communications Act 2003, which applied to communications that a defendant:

1. Sends over a public electronic communications network
2. For the purpose of causing annoyance, inconvenience or needless anxiety to another.
3. And that he knows to be false.¹⁶

We believe the updated offence under the OSA represents an improvement, as it raises the threshold for an offence to have been committed. However, “non-trivial emotional harm” remains broad in scope. A wide range of communications could feasibly fall within this definition and the offence applies to both privately and publicly communicated messages. In addition to the new provisions created by the Act, there are also existing civil remedies, including defamation, harassment and malicious falsehood which are available to individuals who are the subject of the deliberate spreading of false information.

With regard to the regulatory system the OSA establishes the Act also requires service providers to both reduce the risk of and remove content which is illegal from across their platforms. As a result, any disinformation or misinformation that constitutes a hate crime, incitement of violence or the organising and facilitating of rioting would have to be removed under the Act even if the content in question does not result in prosecution of the individual. Any content shared during the Southport riots which constituted one of the aforementioned criminal offences is therefore already subject to

¹⁵Ofcom's approach to implementing the Online Safety Act - Ofcom, 26 October 2024:
<https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/roadmap-to-regulation/>

¹⁶ Harmful Online Communications: The Criminal Offences – Law Commissioner Consultation Paper, pg 146-147:
<https://s3-eu-west-2.amazonaws.com/lawcom-prod-storage-11jsxou24uy7q/uploads/2020/09/Online-Communications-Consultation-Paper-FINAL-with-cover.pdf>

criminal law; introducing additional legislation would unnecessarily complicate this pre-existing framework.

Provisions on misinformation

For truly serious cases of misinformation where the inaccurate information has ramifications for an individual's livelihood, or which causes serious physical or mental harm, the same civil actions that are available for disinformation are likely to apply. Equally, online misinformation which passes the threshold for criminality will be subject to prosecution.

However, debates about the need to tackle online misinformation are generally more concerned with attempts to regulate speech at the societal level, rather than focusing on the effects of misinformation on an individual scale. For instance, Parliament has debated the risks of disinformation and misinformation in the context of general elections,¹⁷ the COVID-19 pandemic and Russian invasion of Ukraine,¹⁸ COP-26 and climate change,¹⁹ the conflict in Israel and Palestine.²⁰ By suggesting that discussions on complex topics are restricted, there is a risk that policymakers and parliamentarians push social media platforms to shepherd debate on topics central to democracy and which frequently engages unpopular, dissenting or fringe views. In practice it would be ineffective to try to moderate widescale 'misinformation' through content moderation, not least, because of the previously discussed definitional problems of determining what is a "mistruth".

Several attempts were made to restrict misinformation directly through the OSA, through amendments which sought to designate various kinds of misinformation as priority content harmful to adults.²¹ However, as outlined above, legislating to restrict misinformation is extremely challenging and open to political abuse. In summing the potential threats to free speech and identifying alternative solutions, the House of Lords Communications and Digital Committee's report into freedom of expression online stated:

¹⁷ HC Deb, Elections: Disinformation, 23 May 2024, vol 750:
<https://hansard.parliament.uk/Commons/2024-05-23/debates/CFE5C4BD-425A-4110-B985-AC178C1DACA2/ElectionsDisinformation>

¹⁸ HL oral question, Counter Disinformation Unit, 18 July 2023, vol 831, col 2200:
<https://hansard.parliament.uk/Lords/2023-07-18/debates/E7127600-965F-4B89-A91A-2B4AE3369B06/CounterDisinformationUnit>

¹⁹ Online Safety Bill Committee (Second sitting), 13 December 2022: [https://hansard.parliament.uk/commons/2022-12-13/debates/9cccea35-f420-499b-8721-ff36c6b1bb30/ONLINESAFETYBILL\(SecondSitting\)](https://hansard.parliament.uk/commons/2022-12-13/debates/9cccea35-f420-499b-8721-ff36c6b1bb30/ONLINESAFETYBILL(SecondSitting))

²⁰ HL oral question, Counter-Disinformation Unit: Israel and Palestine, 24 October 2024, vol 833:
<https://hansard.parliament.uk/Lords/2023-10-24/debates/E41C2A61-E9FC-46D2-8C9F-D349F25AC613/Counter-DisinformationUnitIsraelAndPalestine>

²¹ Online Safety Bill Committee (Tenth sitting), 14 June 2023, col 405:
[https://hansard.parliament.uk/Commons/2022-06-14/debates/be957e3c-244c-48c2-b778-b131667e87af/OnlineSafetyBill\(TenthSitting\)](https://hansard.parliament.uk/Commons/2022-06-14/debates/be957e3c-244c-48c2-b778-b131667e87af/OnlineSafetyBill(TenthSitting)); Baroness Kidron's amendment, After Clause 54, Online Safety Act 2023, Report stage: <https://bills.parliament.uk/bills/3137/stages/17765/amendments/96180>

*"We are not convinced that they are workable or could be implemented without unjustifiable and unprecedented interference in freedom of expression. If a type of content is seriously harmful, it should be defined and criminalised through primary legislation. It would be more effective—and more consistent with the value which has historically been attached to freedom of expression in the UK—to address content which is legal but some may find distressing through strong regulation of the design of platforms, digital citizenship education, and competition regulation."*²²

National Security Online Information Team

In January 2023, Big Brother Watch published our Ministry of Truth report, which revealed how a range of Government units had been secretly tasked with monitoring the online speech of UK citizens.²³ Of these units, of chief concern was the Counter Disinformation Unit ('CDU'). The CDU, which was stood up in 2020 during the COVID-19 pandemic to respond to disinformation and the "inadvertent sharing of false information", is now based in the Department for Science, Innovation and Technology ('DSIT') and has been rebranded as the National Security Online Information Team ('NSOIT'). Since March 2020, the remit of the CDU/NSOIT has expanded dramatically to cover, that we know of, the May 2021 UK local elections, COP-26, the death of Queen Elizabeth II, the NI Assembly elections, and the Russian invasion of Ukraine amongst other issues.

The work of NSOIT involves the monitoring and reporting of the online speech of British citizens, and as such, engages the fundamental rights to freedom of speech and privacy, protected by Article 10 and Article 8 of the European Convention on Human Rights (ECHR) respectively. The European Convention on Human Rights is clear that interferences with these rights are only lawful where they are provided by law, necessary and proportionate.²⁴ The presumption must rest in favour of protecting these rights and interference with them should come as a last resort.

The Government spent approximately £1.5m on contracts for AI company Logically to support the work of the CDU between 2020 and 2023. The contract DCMS awarded to Logically was for "analytical support relating to potentially harmful disinformation online".²⁵ Further, in addition to the contracts Ms Storey outlined, DCMS paid £1.3 million to controversial technology company Faculty Science for a year-long contract

²² Free for all? Freedom of expression in the digital age, House of Lords Communications and Digital Committee, 22 July 2021: <https://committees.parliament.uk/publications/6878/documents/72529/default/>

²³ Ministry of Truth: The secretive government units spying on your speech – Big Brother Watch, January 2023: <https://bigbrotherwatch.org.uk/wp-content/uploads/2023/01/Ministry-of-Truth-Big-Brother-Watch-290123.pdf>

²⁴ The Human Rights Act – Equality and Human Rights Commission: <https://www.equalityhumanrights.com/en/human-rights/human-rights-act>

²⁵ Disinformation Analytical Support – UK Government Contracts Finder, 13 February 2021: <https://www.contractsfinder.service.gov.uk/notice/ecffdb0a-92fd-4e26-ad62-c6c144b4a16b?origin=SearchResults&p=1>

offering “technical support on understanding the threat and mitigation of disinformation” from April 2022.²⁶ Our research, and the documents we obtained via FOI, indicated that the CDU’s main functions were generalised monitoring of social media posts; the production and circulation across government of “mis/disinformation” reports; and flagging posts to social media companies for enforcement.

Our research has uncovered that these “mis/disinformation” reports in fact contain vast amounts of accurate and lawful information. One such report in 2021 titled, “Election Disinformation Report” featured former Green Party MP Caroline Lucas for calling former Prime Minister, Boris Johnson a liar in an interview with the Independent.²⁷ Another disinformation report highlighted a tweet by the established human rights campaigner, Peter Tatchell, in which he shared a video interview criticising the monarchy.²⁸ Another of Tatchell’s tweets discussing the erosion of the right to protest in relation to the state funeral of Queen Elizabeth II was included in a separate disinformation report. Despite these publications being entirely lawful expressions of political opinion, which engaged Tatchell’s free speech rights and could not reasonably be regarded as disinformation, they were monitored and recorded by a contracted third-party and circulated around the government. These are just two examples of many.²⁹ This monitoring of lawful domestic media, campaigners and politicians, as opposed to battling propaganda from hostile entities, raises real questions about the expansion of the Unit’s stated purpose of monitoring national security risk, or “mission creep”.

During the Covid-19 pandemic, the Unit was highly active. The Unit routinely recorded the lawful, truthful speech of individuals who were not conceivably creating or spreading mis or disinformation – including elected politicians, world-leading academics, high-profile journalists and human rights campaigners – in so-called “counter disinformation” reports, as a result of those people criticising the government’s pandemic policies. The topics monitored by the Unit, whether by tracking particular issues or more general “social listening”³⁰ around daily news reports, spanned criticisms of regional support, pandemic preparedness, vaccine

²⁶ Disinformation Technical Services – UK Government Contracts Finder, 27 July 2022: <https://www.contractsfinder.service.gov.uk/notice/5dea9926-0ae8-46ad-8fbc-7bc1fadd2b4c?origin=SearchResults&p=1>

²⁷ Caroline Lucas flagged by disinformation unit over Covid criticism – Peter Walker, the Guardian, 18 June 2023: <https://www.theguardian.com/politics/2023/jun/18/caroline-lucas-flagged-by-disinformation-unit-over-covid-criticism>

²⁸ ‘Snooping’ Claim After Anti-Monarchy Opinions Included in Government Disinformation Reports – Simon Charles, Novara Media, 13 May 2024: <https://novaramedia.com/2024/05/13/snooping-claim-after-anti-monarchy-opinions-included-in-government-disinformation-reports/>

²⁹ For further examples, please see Ministry of Truth: The secretive government units spying on your speech – Big Brother Watch, January 2023: <https://bigbrotherwatch.org.uk/wp-content/uploads/2023/01/Ministry-of-Truth-Big-Brother-Watch-290123.pdf>

³⁰ RAPID RESPONSE UNIT, HMG Daily Social Listening Report – Freedom of Information request, Cabinet Office, 30 November 2021

supplies overseas, vaccine mandates, Covid status certificates, lockdown modelling and protest rights. Those affected include Leader of HM Opposition Sir Keir Starmer KC, Manchester Mayor Andy Burnham, Conservative MP and former minister David Davis, Green MP Caroline Lucas, columnist Peter Hitchens, Dr Alex de Figueiredo (Vaccine Confidence Project, London School of Hygiene and Tropical Medicine), Professor Carl Heneghan (Centre for Evidence-Based Medicine, Oxford University), and many others.³¹

DSIT and NSOIT in particular benefit from 'trusted flagger' status with many social media platforms, which means that any material that the Government reports is prioritised for review and potential action. There is likely to be additional pressure on a platform to censor the flagged material in these circumstances, given that it comes from the state. This is of particular concern where the material which cannot be considered disinformation, let alone illegal content – as was the case in the examples of content covered in misinformation reports. Any relationship that exists between the Government and social media companies should be both rights-respecting and transparent. At present, the Government is able to flag content that it considers to be objectionable and in breach of platforms' terms of service at its own discretion, undermining the universal application of the right to free expression.

The social media company X, formerly Twitter, revealed that 58% of flags it received from the CDU between November 2020 – February 2024 did not violate their terms of service.³² At the time, Twitter's terms of service were incredibly broad and were even expanded publicly on 1 April 2020 to include harmful speech "that goes directly against guidance from authoritative sources of global and local public health information"³³ – in other words misinformation. This means that the majority of flags made by the government to Twitter during this period did not in fact breach the platform's terms of service, nor did it contradict authoritative public health guidance. This suggests that NSOIT plays a role in encouraging the adoption of an overly censorial approach to content moderation by social media companies. Indeed, between 2021 and 2023, when Big Brother Watch published our Ministry of Truth report, the number of times that the government flagged speech to social media platforms plummeted by 95%.³⁴ We remain concerned that lawful speech continues to be flagged and removed through these inscrutable extrajudicial requests by the government.

31 Ministry of Truth: The secretive government units spying on your speech – Big Brother Watch, January 2023, pg 38, 12, 36: <https://bigbrotherwatch.org.uk/wp-content/uploads/2023/01/Ministry-of-Truth-Big-Brother-Watch-290123.pdf>; Secret government unit collected Telegram posts about Covid policy critic – the Telegraph, 7 June 2023: <https://www.telegraph.co.uk/politics/2023/06/07/covid-19-carl-heneghan-counter-disinformation-unit-telegram/>

32 Twitter refused majority of removal requests from Covid spying unit- the Telegraph, 10 June 2023: <https://www.telegraph.co.uk/news/2023/06/10/twitter-refused-removal-requests-covid-spy-unit/>

33 Coronavirus: Staying safe and informed on Twitter – X, 1 April 2020: https://blog.twitter.com/en_us/topics/company/2020/covid-19#definition

34 Government censorship unit sees 95% fall in 'flagged' online content – Mark Johnson, UnHerd, 1 June 2024: <https://unherd.com/newsroom/government-censorship-unit-sees-95-fall-in-flagged-online-content/>

New NSOIT policy

There is limited public information about NSOIT's operations. Ministers have refused to provide a comprehensive list of the topics that are being monitored, and will not provide the number of civil servants working as part of the Unit.³⁵ After freedom of information requests and an intervention from the Information Commissioner's Office, the Government provided us with NSOIT's policy in April 2024. Tellingly, there are almost no references to national security in the policy document. Instead, the policy document states that NSOIT's function is "to build a picture of mis/disinformation threats". NSOIT's broad remit also includes issues relating to 'public safety'. Renaming the CDU appears to be an attempt to deflect criticism of the Unit by claiming it has a national security remit, and focuses on foreign threats rather than domestic speech. In its policy document however, it is clear that the unit's remit for analysis is virtually limitless. While it states that any "collection activity" must be linked to a government policy goal, this provides almost unlimited scope. NSOIT policy states it "does not seek to capture genuine political debate". The designation of some political debate as 'genuine' and some as not is concerning and open to abuse.

NSOIT is still required to flag content to social media platforms, if the content is considered to be harmful, violates the platform's Terms of Service, or represents an emerging threat. These categories offer little protections from lawful speech being by flagged by NSOIT analysts for removal. Indeed, social media platforms' Terms of Service restrict speech far beyond the law, and by working with platforms to have content removed, this policy undermines the government's obligations to uphold freedom of expression.

Parliamentary response

In December 2022, the Intelligence and Security Committee ('ISC'), complained of an "erosion of oversight" and claimed that the Government was "refusing" to expand the Committee's remit to include the CDU.³⁶ In the report the Committee expressed concern that "intelligence activities" are being increasingly devolved to policy departments, such as the Department for Culture, Media and Sport (now DSIT). The ISC recommended that the CDU/NSOIT, be added to its remit to ensure proper scrutiny of intelligence activities conducted by ordinary Whitehall departments.

In March 2024, Parliament's Culture, Media and Sport Committee published a report, *Trusted Voices*, expressing a "lack of transparency and accountability of the CDU and

³⁵ Written question: National Security Online Information Team, UIN 3443, tabled on 30 August 2024: <https://questions-statements.parliament.uk/written-questions/detail/2024-08-30/3443>

³⁶ Press Release: Annual Report 2021-22 – Intelligence and Security Committee, 13 December 2022: <https://isc.independent.gov.uk/wp-content/uploads/2022/12/ISC-Annual-Report-2021%E2%80%932022-Press-release.pdf>

the appropriateness of its reach.”³⁷ The report recommended that “the Government commission and lay before Parliament an independent review of the activities and strategy of Counter Disinformation Unit within the next 12 months.” Additionally, in September 2023, a group of cross-party MPs, including Caroline Lucas, David Davis and Bell Ribeiro-Addy, called for a review of the unit.³⁸

We are concerned that the new Government has responded to the recommendation of the Culture, Media and Sport Committee by ruling out a review of NSOIT, instead stating that ministerial oversight of the Unit is sufficient.³⁹ We believe this is an inadequate level of scrutiny and oversight, given the impact NSOIT has on freedom of expression online and the absence of transparency in which it has operated. The concerns raised by the Culture, Media and Sport Committee over the operation and governance of the CDU have not been allayed by changing the name of the Unit to NSOIT. In order to protect the public’s rights to free speech and privacy and ensure this body operates transparently, lawfully and in an accountable manner, it is incumbent on the Government to urgently commission a review of NSOIT.

Conclusion

Concerns over the spread of mis/disinformation are legitimate in the internet age, where content can be targeted at individuals like never before and shared rapidly. We believe that the Government should focus on the business models of social media platforms that profit of mass data collection and aim to keep users on their sites for as long as possible by feeding them incendiary content.

Rather than attempting to act as the arbiter of truth and legislate for the removal of individual pieces of content, the Government should focus on promoting media literacy, algorithmic transparency and expanding user controls.

Further, we reiterate our recommendation that the Government instigate an independent review of the work of the National Security Online Information Team.

37 Trusted voices – Culture, Media and Sport Committee, Sixth Report of Session 2023–24, HC 175, 26 March 2024: <https://committees.parliament.uk/publications/44146/documents/219482/default/>

38 Disinformation unit ‘a stain on democracy’ that must be axed, say MPs – the Telegraph, 11 September 2023: <https://www.telegraph.co.uk/politics/2023/09/11/cdu-suppressing-free-speech-covid-government-opposition/>

39 Trusted voices: Government response – Culture, Media and Sport Committee, Second Special Report of Session 2024–25, HC 292, 4 November 2024: <https://committees.parliament.uk/publications/45517/documents/225336/default/>