Mark Bunting

Keeping Consumers Safe Online

Legislating for platform accountability for online content

July 2018

COMMUNICATIONS
CHAMBERS

"The internet is growing in importance around the world in people's lives and I think that it is inevitable that there will need to be some regulation"

Mark Zuckerberg, testimony to US House of Representatives Energy and Commerce Committee, 11 April 2018

"We need to now take a more active view in policing the ecosystem…At the end of the day, this is going to be something where people will measure us by our results."

Mark Zuckerberg, testimony to US Senate Committees on the Judiciary and Commerce, Science and Transportation, 10 April 2018

"If you create technology that changes the world, the world is going to want to govern [and] regulate you. You have to come to terms with that."

Brad Smith, Microsoft, 29 May 2018

About the Author

Mark Bunting is a Partner at Communications Chambers, having previously held senior strategy roles at the BBC and Ofcom. He has worked on the implications of the growth of platforms for content regulation as a visiting associate at the Oxford Internet Institute,[1] and coauthored a Communications Chambers report, funded by Apple, on principles for platform regulation.[2]

---

[1] Mark Bunting, *From editorial obligation to procedural accountability: new policy approaches to online content in the era of information intermediaries*, May 2018

[2] Brian Williamson & Mark Bunting, *Reconciling private market governance and law: A policy primer for digital platforms*, May 2018

Contents

[blank page]

1. Executive summary

Platforms regulate online content, but lack oversight

In a world of effectively infinite information, online platforms play a vital role in selecting, organising, ranking, recommending and suppressing content and content providers.

Platforms (or 'online content intermediaries') now find themselves on the frontline of online content regulation: an inevitable consequence of users' ability to post content that ranges from illegal and seriously harmful, to legal but socially unwanted. Their rules are explicit, in community standards and terms of use, but also in the implicit and sometimes unintended effects of personalisation algorithms.

There is currently no systematic means of assessing the impact of platform's content policies, algorithms and decisions, nor of holding intermediaries to account. Platforms are not above the law, but European law limits their liability to conditions in which they have 'actual knowledge' of illegal content.

A consensus is growing that further intervention is needed to address platforms' role in governing online content, given its importance to the public interest in a host of areas. However, society's expectations for this broader role have not yet been codified, either with respect to desired outcomes or good governance procedures.

Arguably, this is the single biggest gap in Internet regulation, although it could be addressed relatively easily. The state plays a role in setting standards in most other information markets, recognising the social harms and benefits of certain kinds of content. It could do the same here. Doing so need not conflict with European law, which anticipates

states may apply a duty of care on intermediaries in relation to illegal activity.

Intermediaries' policies vary widely. Some variation is appropriate, as users' expectations also vary, but consumers have a right to know what to expect from platforms. Today's fragmented (and not always well-advertised or effectively enforced) standards and processes may not help. Some platforms work closely with governments to address potential content harms, others do not.

A statutory framework for intermediary accountability

The UK Government has committed to bring forward a White Paper on online harms and safety, including a Code for Practice and Transparency Reports. The White Paper could establish a wider statutory framework for platform accountability for online content.

Such a framework would aim to:

- Clarify what consumers can expect from intermediaries, in their handling of harmful and illegal content;
- Ensure intermediaries' governance of online content is proportionate and accountable, and takes a fair and responsible approach to balancing rights;
- In achieving these goals, recognise differences between intermediaries of varying size and different business models, and the need for regulatory certainty and an outcomes-based approach.

Core components of the framework would be:

- A Code of Practice for content intermediaries, defining broad content standards and procedural expectations

- A List of intermediaries in scope for different tiers of obligation
- Incentives and sanctions to encourage intermediaries to adhere to the Code
- An independent oversight body, tasked with maintaining the Code and the List, requesting certain information from intermediaries, promoting consumers' and intermediaries' rights and responsibilities, and reporting publicly on the effectiveness of platforms' content policies.

Many elements of this framework are contained in existing law and regulation, but they are not consistently brought together.

The envisaged Code would be broad and flexible enough to adapt to new concerns and platforms. Its requirements would be proportionate to evidence of harm, with the priority on illegal and seriously harmful content, and lower expectations for legal-but-harmful material. It would also differentiate on size, with reduced or no requirements for smaller platforms. The baseline requirements of intermediaries above a de minimis size would be to notify the oversight body, contribute to its costs of operation, and provide information or carry out a harm assessment in response to a specific, evidence-based and reasonable request.

The oversight body could be industry-led, if intermediaries can form an independent organisation with industry and Government support, able to make binding decisions, with a backstop regulator fulfilling a role in this co-regulatory model. The Advertising Standards Authority offers a precedent.

Or oversight could be provided by a statutory body (either an existing institution such as Ofcom, or a new body), in which case it should be funded by industry, as Ofcom is today.

Benefits and risks

In this model, responsibility for actual content regulation – policy development, notification and appeals, use of automated detection tools, human moderation – continues to sit with intermediaries themselves, who are best placed to govern platforms in users' interests.

The purpose of these proposals is to provide better oversight of that activity, and thereby replace 'regulation by outrage' with a more effective and proportionate approach.

All stakeholders could benefit from such a model, including intermediaries, who would have greater clarity about what is expected of them, the legitimacy that comes from external scrutiny and validation, and defence against unreasonable or unevidenced requests.

Oversight needs to be cautious, and limited in statute, to mitigate potential risks to openness, innovation, competition and free speech. We believe this model does not require changes to intermediaries' liabilities, although in the longer-term a review of the E-Commerce Directive may be appropriate.

Recommendations

Government should include an accountability framework for online content intermediaries in the planned White Paper on online harms and safety. This should make provision for a Code of Practice, an oversight body and incentives and sanctions.

Industry should consider the potential to form a co-regulatory body to provide independent oversight of intermediaries' content policies, with buy-in from most platforms with significant numbers of UK users.

Government should develop options for a statutory oversight body, in case the industry option does not make sufficient progress within a reasonable time period.

2. About this report

Platforms, rules and the 'wild west'

The Internet is often described as an unregulated haven for harmful and illegal content. Concerns are raised about extremist propaganda, hate speech, inappropriate videos targeting children, online abuse and harassment, fraud, intellectual property infringement, fake news, child sexual imagery and many other issues.

None of these issues are new, but over the past couple of years we have learnt more about the efforts of major platforms, or 'online content intermediaries', to try to limit access to such content.

In fact, there is a great deal of content regulation online – the biggest platforms use automated technology to block or remove huge quantities of material and employ thousands of human moderators. Often this has involved some form of government encouragement or involvement, whether through bilateral discussions or through cross-industry and multi-stakeholder initiatives such as the Internet Watch Foundation or the Global Internet Forum to Counter Terrorism. However, the effectiveness and impact of these activities are hard to judge, and beyond the larger firms, efforts are more variable.

Section 3 describes the challenges and opportunities these developments pose for policy-makers. Intermediaries face intense pressure to "do more," in turn raising concerns about free speech and the centralised power of big platforms. Even the wild west had rules; today's online sheriffs are private firms whose policies, decision processes and enforcement actions can be opaque and subject to little external accountability. The debate about whether intermediaries are 'platforms or publishers' may have helped mask a more important question: how can intermediaries' role in governing access to online content be more effectively and accountably harnessed in the public interest?

The policy comeback

Section 4 assesses regulatory responses to this question. Governments around the world are considering options; in Europe, this is putting strain on the legal framework established in the E-Commerce Directive (despite strenuous efforts by the EU's own policy proposals to avoid conflict with it). The UK Government has committed to bring forward online safety legislation as part of its Digital Charter.[3]

Two varieties of regulatory requirement can be distinguished: proactive obligations, which direct intermediaries to block or remove specified content prior to notification; and procedural requirements, which set standards for the processes intermediaries use to manage and account for the content they host.

Accountable governance

Section 5 describes a possible blueprint for an accountability framework for online content intermediaries. It suggests that the Government should use the opportunity of legislation to:

- make provision for a Code of Practice that defines broad content standards and procedural expectations of intermediaries, without prescribing particular actions or technological solutions;

---

[3] HM Government, _Government response to the Internet Safety Strategy Green Paper_, May 2018

COMMUNICATIONS
CHAMBERS

- define the scope of intermediary accountability and set thresholds for regulatory intervention, to ensure that obligations are proportionate to harm and smaller intermediaries do not face an excessive burden;
- provide for an oversight body to oversee harm assessments and validate the effectiveness of intermediaries' policies;
- provide the statutory basis for incentives and sanctions required to encourage compliance with the Code of Practice; and
- establish due process, including intermediaries' and users' rights of appeal and recourse.

This paper does not describe detailed rules or processes, but outlines the main components of this accountability framework and suggest how it could be developed further. Section 6 assesses what legislation would be needed to provide for the framework, and considers the pros and cons of oversight by industry or statutory bodies.

Implementation

Section 7 describes the next steps required to take this proposal forward. We consider that it could be implemented consistently with the E-Commerce Directive, although in the longer term a review of some aspects of the Directive may be appropriate. In the short-term, more work is needed, with intermediaries and other stakeholders, on detailed questions of definition, scope, thresholds for intervention and the oversight role, to inform the Government's planned White Paper.

About Communications Chambers

Communications Chambers is an advisory firm specialising in technology, media and telecoms policy and strategy. Its members have worked at senior levels in industry, regulators and government.

Mark Bunting
mark@commcham.com

3. Why regulate online content platforms?

The rise of platforms

The Internet has transformed the supply and consumption of content by dramatically lowering the costs of distribution. Providers of information and content services can access a global market at very low cost, bringing dynamism, innovation and competition to markets that were historically often dominated by entrenched firms.

Publishing, broadcasting, journalism, business intelligence and academic publishing are amongst the sectors that have been transformed by online competition; consumers have benefited from new entrants, new content formats and more choice about how, where and when they consume. Advertisers have new ways to reach audiences, with greater ability to target both demographic and behavioural characteristics.

As barriers to entry fell, the supply of content available to consumers became infinite, for all practical purposes. Intermediaries emerged to help consumers, content providers, and advertisers come together, and to help their users find the content that best meets their needs. This is an essential function of the digital economy; as Parker, van Alstyne and Choudary put it, "frictionless entry must be balanced by effective curation."[4]

Directories, search engines, aggregators, app stores and social media all, in different ways, provide this curation function. Some do so on a vast scale: every day over 500,000 hours of content are uploaded to YouTube, billions of items of content are shared on Facebook, and 500m tweets are posted.

This paper focuses on 'online content intermediaries' who play an active role in managing the content choices available on their platforms (for more on terminology, see Box 1). For our purposes, online content intermediaries have three defining features:

- They operate open marketplaces (or 'multi-sided markets') that create value by enabling direct interaction between suppliers and consumers of information and content;
- They play an active role in matching content to users, by selecting, sorting, ranking, recommending or suppressing content;
- They earn revenue by taking a share of the value their platforms create, including by hosting advertising around content, or by earning commission on transactions consummated on their platforms.

Intermediaries give the lie to the view that online content is impossible to regulate. They set rules ('Community Standards') about what

---

Box 1. Note about terminology

The term 'platforms' is often used as a catch-all for intermediaries, and we use the two terms broadly interchangeably here. However, 'platform' does not appear in the relevant European legislation. We use 'online content intermediary' more specifically to describe a subset of 'hosting' providers, in the language of the E-Commerce Directive. It has similarities to the concept of an 'active' host, which originated in a more limited sense in European case law (*L'Oreal v eBay*), and has been developed further in recital 38 of the draft *European Copyright Directive*.

---

[44] Geoffrey Parker, Marshall van Alstyne, Sangeet Choudary, *Platform Revolution: How Networked Markets are Transforming the Economy – and How to Make Them Work for You*, W.W. Norton, 2016, p26

kinds of content are permissible on their platforms, and enforce them with both automated tools and human moderation. They may also promote some kinds of content over others,[5] and shape providers' business models and commercial incentives, for example by deciding whether to allow third parties to offer subscriptions via their sites.[6]

As Communications Chambers has said in a previous report, "a complaint, or complement insofar as innovation is concerned, is that the new players are "lawless", but a deeper concern may be that they are "law makers" – in terms of code versus law."[7]

This governance role should be policy-makers' focus. The sterile debate about whether intermediaries are 'platforms or publishers' needs to be abandoned (see Box 2).

While most content hosted by intermediaries is innocent,[8] challenges arise when providers and consumers of harmful and illegal content exploit the openness of online platforms. Access enables abuse; 'bad actors' are the inevitable counterpart of the unprecedented freedom of expression platforms offer.

Many sites have developed tools to provide 'safe' content environments, in their users' interests, including automated detection tools and user controls. Facebook was working on tools to automatically flag pornography, hate speech and bullying as early as 2015.[9] Its recent Transparency Report disclosed that it had removed 583m fake accounts in the first quarter of 2018, and 21m pieces of content featuring sex or nudity.[10] Google says that 81% of extremist content it takes down are removed by automated tools.[11] Instagram has designed filters and comments screening tools to suppress abusive content and spam.[12]

However, perfection is impossible. Determined suppliers and consumers of illegal content will always find a way round platform rules and algorithms.

Box 2. Platforms or publishers? The risk of definitional distraction

The longstanding debate about whether intermediaries are 'platforms or publishers' has not helped illuminate the challenges of Internet regulation. Content intermediaries include a very wide range of businesses, from search engines to app stores, and shoehorning them into legal frameworks from another technological era is a mistake. In section 4, we argue that policies that treat platforms as publishers in practice merely delegate legal decision-making to intermediaries, with limited oversight and accountability, and may advantage larger firms over competitors.

We need a new definition, that focuses on the particular role intermediaries play in online information markets. This can be thought of as a form of market governance: the design, implementation and enforcement of rules concerning the free flow of information and content exchange. Unlike 'mere conduits', this governance function means content intermediaries cannot be neutral between different types of content. The content and effects of their community standards, personalisation algorithms and commercial policies will inevitably be controversial.

---

[5] Facebook, *Helping Ensure News on Facebook Is From Trusted Sources*, 19 January 2018
[6] Facebook, *Testing Subscriptions Support in Instant Articles*, 19 October 2017
[7] Williamson & Bunting, *supra* note 2
[8] Facebook says that of every 10,000 content views on its platform, 7-9 violate its rules on adult nudity and sexual activity
[9] The Verge, *Facebook is patenting a tool that could help automate removal of fake news*, 7 December 2016
[10] Facebook, *Community Standards Enforcement Preliminary Report*, May 2018
[11] Google, *YouTube Community Guidelines enforcement*, Oct-Dec 2017
[12] Wired, *Instagram's CEO Wants to Clean Up the Internet*, August 2017

Different intermediaries take very different approaches to online safety. Some variation is important and appropriate, as expectations of different platforms differ. But it is usually possible to define shared broad standards. For example, the Google Play Store, Reddit and Musical.ly are very different platforms, whose users expect different standards. But all have policies on hate speech, even though they implement them in very different ways.

Why regulation is needed

Many firms play a market governance role without regulation having historically been considered necessary, even if incentives may not be aligned with policy goals. Credit card companies, operating system providers and food processing firms may all be thought of as platforms that play a role in deciding who can enter markets, and on what terms.

But there are five reasons regulation may be considered appropriate for online content intermediaries.

First, information goods play a special role in society, supporting democratic engagement, promoting the transmission of ideas, building community identities and enabling economic empowerment.[13] Fundamental rights must be balanced: freedom of expression, respect for privacy, dignity and non-discrimination, protection of intellectual property and the right to conduct business.[14] Both positive and negative externalities in content markets are policy concerns, and a range of legal and regulatory requirements address them in other areas (see Box 3).

Online platforms are not immune from these concerns, even if specific standards do not read across. As the Council of Europe says:

> "The power of such intermediaries as protagonists of online expression makes it imperative to clarify their role and impact on human rights, as well as their corresponding duties and responsibilities, including as regards the risk of misuse by criminals of the intermediaries' services and infrastructure… States are confronted with the complex challenge of regulating an environment in which private parties fulfil a crucial role in providing services with significant public service value."[15]

Box 3. Law and regulation in other content markets

Most content markets are regulated, either statutorily or voluntarily (although usually with the background threat of statutory regulation). For instance:

- publishers are liable for damage or injury arising from defamatory or libellous content, or copyright infringement;
- the press is self-regulated, in line with the Editors' Code of Practice and (for those bodies signed up to it) the IMPRESS Standards Code;
- broadcasting is subject to a range of statutory regulation by Ofcom, both to reduce the availability of harmful and illegal content, and to promote positive objectives such as freedom of expression and media diversity; and
- advertising content is regulated by the Advertising Standards Authority (an industry body, with Ofcom as a statutory backstop).

---

[13] Joseph Stiglitz, *Knowledge as a global public good*, 1999; Ofcom, first PSB Review
[14] EU *Charter of Fundamental Rights*, 2000
[15] Council of Europe, *Recommendation on the roles and responsibilities of internet intermediaries*, March 2018

COMMUNICATIONS
CHAMBERS

Second, the cost or disadvantage of exiting certain platforms may reduce choice and effectively lock some users in. In theory, if a user does not like Facebook's content policy or trolling on Twitter, they can go elsewhere, but they may find few of their friends or intended recipients there.[16]

Broadcasting regulation was historically posited in part on the limited number of channels being available to audiences. Online, bottlenecks have not disappeared, but they have shifted to the discovery of content, making the rules by which intermediaries control discovery a matter of legitimate public interest.[17]

Third, the commercially efficient response to harmful or illegal content is not removal, but personalisation. Personalisation tools and algorithms, in theory, allow intermediaries to ensure each individual gets the particular content that maximises their value from the service. But the greater the degree of personalisation, the more opportunities exist for providers and consumers of harmful content to find each other, often largely unseen by other platform users.

Fourth, notwithstanding these incentives, many intermediaries already actively regulate online content, including to meet policy goals. But the impact of this action is unclear, and accountability is limited. Transparency standards vary, making it hard to compare effectiveness across the industry. It is rare that companies report on all relevant metrics, for example the amount of legal content wrongly flagged by automated tools, as well as content

taken down. Smaller platforms may lack the resources, but not necessarily the challenges, faced by bigger firms.

Finally, and linked to all the preceding points, intermediaries may need external oversight to secure legitimacy and consumer trust. Doteveryone's research found that consumers are confused about what rules, if any, govern online services, and who to go to for advice or recourse.[18] It also shows that many people doubt the Internet's positive impact on society. Edelman's Trust Barometer found increased doubt that technology companies are adequately transparent in 2018, and that while overall trust in technology remains strong, it is declining in many countries, including trust in platforms.[19]

Conclusion: purposes of regulation

The overarching policy goal is to establish new norms, about acceptable behaviours online, the rights and responsibilities of different users, and the role of intermediaries in balancing those rights. Regulation is needed, if markets face systemic problems and social costs that are not fully internalised.

The arguments presented here suggest that such problems are inherent in online content markets. Regulation should be considered, with the specific purpose of:

- clarifying what consumers should be able to expect from intermediaries, in their handling of harmful and illegal content;
- ensuring intermediaries' governance of online content is proportionate and

---

[16] Facebook and Twitter may be eclipsed by rivals over time, just as MySpace and Usenet were before them, but the powerful network effects characteristic of online platforms raise switching costs, and mean the online content market is likely to continue to be dominated by a small number of large, centralised intermediaries

[17] Natali Helberger, Katharina Kleinen-von Königslöw, Rob van der Noll, *Regulating the new information intermediaries as gatekeepers of information diversity*, Info 17(6), 2015

[18] Doteveryone, *People, Power and Technology: The 2018 Digital Attitudes Report*, February 2018

[19] Edelman Trust Barometer, *Trust in Technology*, March 2018

accountable, and takes a fair and responsible approach to balancing rights;

- in pursuing these objectives, recognising differences between intermediaries of varying size and different business models, and the need for regulatory certainty and an outcomes-based approach.

In the next section, we assess whether existing legal and policy frameworks have achieved these purposes, and whether current policy developments offer a solution.

4. UK, European and global policy responses

Current law and regulation in the UK

There have been two major pieces of digital legislation in the UK, the Digital Economy Acts of 2010 and 2017. However neither involved a comprehensive review of content regulation, the current frameworks for which date back to the 2003 Communications Act (which did not mention online content or platforms) and the 2000 European E-Commerce Directive (ECD).[20]

The ECD defines three types of intermediary, 'mere conduits', 'caching providers' and 'hosts' – the latter being the category of interest here. Intermediaries that host illegal content incur legal liability for it *only* if they have 'actual knowledge' of it and do not act 'expeditiously' to remove or block it.

The definition of host is wide and operates across many different business models. Subsequent case law has introduced the concept of 'active' hosts, who are liable for content they treat 'non-neutrally'.[21]

The ECD has led to 'notice and takedown' being the main mechanism by which third parties – users, content providers, governments – solicit action by intermediaries. YouTube receives 3.1m flags of inappropriate videos per month.[22] Facebook says it receives 280,000 reports per month of hate speech;[23] Google gets about 10,000 requests per month, asking for roughly 40,000 URLs to be delisted under the 'right to be forgotten'.[24]

Liability limitations are an important feature of the digital economy, helping to preserve the openness of online platforms, low barriers to entry and capacity for innovation, while also providing a means of incentivising intermediaries to act on identified illegal content.

But it does have disadvantages. On the one hand, it means intermediaries cannot be held responsible even for systematic, predictable and potentially intentional abuse enabled by their platforms. Citron and Wittes give examples of sites created specifically to allow distribution of revenge porn, a gossip site that urged users to send in 'dirt', a platform for sex trade adverts whose policies and architecture were designed to make detection of sex trafficking difficult, and a site for 'talking to strangers' with no age verification tools.[25]

Platforms are often also criticised for being slow to respond to reports of illegal content; a German study found in 2017 that Facebook and Twitter had failed to meet a national target of removing 70% of hate speech within 24 hours of notification.[26] Advertisers complain that it is impossible for them to reliably prevent their adverts being placed next to inappropriate content.[27]

On the other hand, some intermediaries now take discretionary steps on harmful and illegal content that go significantly beyond their legal liabilities (see section 3). This creates some legal risk to them, as it may not be clear when

---

[20] https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32000L0031

[21] Cyberleagle, *The Electronic Commerce Directive - a phantom demon?*, April 2018

[22] Google, *supra* note 11

[23] Karim Palant, UK Public Policy Manager, Facebook, at Westminster Media Forum conference, 16 January 2018

[24] Google, *Search removals under European privacy law*, Oct-Dec 2017

[25] Danielle Keats Citron and Benjamin Wittes, *The Internet Will Not Break: Denying Bad Samaritans § 230 Immunity*, Fordham Law Review vol 86 issue 2, 2017

[26] New York Times, *Delete Hate Speech or Pay Up, Germany Tells Social Media Companies*, 30 June 2017

[27] Digiday, *The new most important role at agencies: Brand-safety officer*, 11 April 2018

these steps give rise to 'actual knowledge' that would trigger liability.

More importantly, there is currently no way of assessing the impact and effectiveness of these activities, either with respect to takedown of illegal material or inadvertent blocking of legal content. Evaluations are generally conducted by intermediaries themselves, who choose what to measure and disclose. While the many transparency reports provided by the likes of Google, Facebook, Twitter and others are useful, they do not represent a comprehensive assessment of the impact of their content governance activities.

'Regulation by outrage'

Outrage, campaigning and lobbying have stepped into this gap. The pattern has become familiar: a problem is identified, media coverage intensifies, political pressure is applied and threats of regulation are issued. Intermediaries respond, with *mea culpa*s and promises to do better. Initiatives are launched, either by individual companies, or at industry level, and with varying degrees of involvement by regulators and Government. All parties are able to claim 'something has been done', but exactly what, and to what effect, may remain unclear.

'Regulation by outrage' puts issues on the political and corporate agenda, but leaves consumers none the wiser about the true risks of online content nor what they have a right to expect from intermediaries. It may aggravate mistrust of intermediaries, while doing little to empower consumers with accurate and impartial information.

The sustainability of this state of affairs is now under question. As David Anderson QC recently put it, "the current state of content regulation is rudimentary, fractured and – it seems fair to assume – transitional... Outside observers note the sub-optimal nature of a system under which content rules are devised in an ad hoc manner by private companies, under pressure from political and commercial interests and without public debate or visibility."[28]

It seems unlikely that a self-regulatory approach will now be seen as a sufficient response to these challenges; any industry-led response risks lacking trust and legitimacy. As Ofcom has noted in another context:

> "in the absence of alignment between the interests of the industry and the public interest, self-regulatory regimes are unlikely to prove effective when confronted by circumstances which present a tension between the public interest and the corporate interests of industry players."[29]

Even if such regimes can be established, it may be difficult to assess their effectiveness without, at minimum, independent scrutiny and reporting.

New policy proposals

Governments around the world are considering how to engage intermediaries' role in governing online content:

- In Europe, proposals for the revised Copyright Directive and Audiovisual Media Services (AVMS) Directive introduce new intermediary duties. The European

---

[28] Counsel Magazine, *Who governs the Internet?*, April 2018
[29] Ofcom, *Submission to the Leveson Inquiry on the future of press regulation: A response to Lord Justice Leveson's request*, April 2012

Commission, Council and Parliament have called for platforms to strengthen measures to tackle illegal and harmful content, including automated detection tools;[30]

- Germany has passed the Network Enforcement Act, or 'NetzDG', introducing large fines for platforms that persistently fail to remove hate speech within 24 hours of notification;[31]
- France and Italy introduced plans to limit online disinformation, including action by intermediaries;[32]
- The US has passed legislation making websites liable for knowingly publishing content designed to facilitate sex trafficking, which amends the safe harbours established by the 1996 Communications Decency Act;[33]
- The Australian Parliament has proposed a duty of care for social media platforms, and "significant financial penalties" for failure to address cyber-bullying[34] (Australia has long-standing Internet Industry Codes of Practice, overseen by Communications Alliance, an industry body[35]
- Singapore has announced plans to legislate for 'fake news'; and
- Sri Lanka and Papua Guinea both temporarily blocked access to Facebook[36]

Setting aside site-blocking, these proposals contain two main approaches.

Proactive obligations

Some proposals require particular actions by platforms to address harmful or illegal content, including proactive measures to remove it prior to notification.

A drift to proactive obligations is particularly apparent in European policy debates. Draft revisions to the Copyright Directive require intermediaries that host 'significant amounts' of user-uploaded content to prevent access to copyrighted works. The draft AVMS Directive requires video-sharing platforms to put in place measures to protect children, and protect all citizens from content that incites violence or hatred.

Such obligations are only meaningful if they come with means of assessing whether their objectives have been met. But measures of success, and penalties for not achieving them, are rarely defined. They may result in legal material being inadvertently blocked. They are likely to be easier for larger, better resourced platforms to implement, locking in the power of incumbents and increasing costs of innovative entry. Some critics have argued that these proposals appear inconsistent with the ECD's prohibition on 'general obligations to monitor' content hosted by platforms.[37]

Most importantly, proactive obligations do not address the accountability gap described above, or promote a fair and responsible balancing of rights. Instead they give

---

[30] European Commission, *Recommendation on measures to effectively tackle illegal content online*, March 2018

[31] DW.com, *Germany implements new internet hate speech crackdown*, January 2018

[32] The Atlantic, *Italy Scrambles to Fight Misinformation Ahead of Its Elections*, 24 February 2018

[33] Engadget, *Trump signs controversial FOSTA-SESTA bill into law*, 11 April 2018

[34] Sidney Morning Herald, *Stricter regulation and financial penalties on the cards for social media platforms*, 29 March 2018

[35] *http://www.commsalliance.com.au/Documents/all/codes/content_services*, *http://www.commsalliance.com.au/Documents/all/codes/internet_and_mobile_codes*

[36] Buzzfeed, *"We had to stop Facebook": when anti-Muslim violence goes viral*, 7 April 2018; Telegraph, *Papua New Guinea blocks Facebook for a month to assess impact on society*, 29 May 2018

[37] Giancarlo Frosio, *The Death of 'No Monitoring Obligations': A Story of Untameable Monsters*, Journal of Intellectual Property, Information Technology and Electronic Commerce Law, vol 8 issue 3, 2017

platforms great power to decide what is legal and what is not, often using automated tools, with little oversight of how they use that power or its full impact on consumer and citizen interests.

Critics have argued that intermediaries are likely to 'over-compensate' in response to proactive obligations, taking down too much content to minimise regulatory risk.[38] The less clearly-defined the targeted content is, and the more important context is to its interpretation, the more likely platforms are get the judgement wrong. For example, YouTube deleted video evidence relating to potential war crimes in Syria as part of its efforts to remove extremist content.[39]

Procedural obligations

Procedural requirements relate to the processes by which platforms manage the content they host. For example, the draft AVMS Directive sets out a range of activities that might be expected, in protecting users from harmful content: terms and conditions, user flagging, user rating, age verification, parental controls.

Procedural obligations shift the focus from *what* platforms do to *how* they govern content markets. This is helpful, given the downsides of proactive obligations, and the difficulty of drawing a boundary between legal and illegal, distasteful vs harmful.

But procedural requirements can still be too specific. Given the diversity of platform environments, it is unlikely that any particular set of procedures is right for all intermediaries. Platforms change rapidly, so procedural requirements risk becoming out-dated.

General procedural standards, based on universal principles – the need to clearly define problems, gather evidence, identify and evaluate solutions, resolve disputes – may be more appropriate than directive regulation.

Conclusion: the policy gap

Neither the existing legal framework, nor self-regulation, nor the shift to proactive obligations, especially in European regulation, provide a comprehensive solution to the regulatory purposes described in section 3. The ECD has created an industry in 'notice and takedown' mechanisms as the main form of platform oversight, but leaves platforms' much wider role in regulating online content largely to their discretion.

Similarly, none of the existing initiatives described here provide broad standards that could apply across a range of content types and intermediaries, and that allow flexibility for different intermediaries to take different approaches to mitigating harm, consistent with a general set of good governance principles.

Arguably, this gap is the most significant in Internet law and regulation. It needs to be clearly distinguished and tackled separately from concerns about consumer protection, use of data and privacy (where GDPR provides a response, and needs time to bed down).

Some have argued that the underlying problem is with the market dominance of certain intermediaries. However, while competition problems may exist, competition law remedies are not generally well suited to addressing content-related externalities. Promoting competition may simply fragment

---

[38] Frank La Rue, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, UN General Assembly Human Rights Council, 2011

[39] Sky News, *YouTube 'deleted Syrian war crime evidence'*, 15 August 2017

content-related problems, making them harder rather than easier to deal with.

Regulation requires specific solutions to well-defined problems. As Ofcom puts it:

> "Effective regulation requires a clear definition of the services that are to be regulated, a specific account of the potential harm to be addressed, and

hence a clear rationale for the specific regulation."[40]

The ingredients of a comprehensive framework for the accountability of online content intermediaries exist, but they are not brought together. In the next section we discuss what such a framework would involve.

---

[40] Ofcom, *Written evidence to the House of Lords European Union Committee Inquiry on Online Platforms and the Digital Single Market*, October 2015

5. Design for an accountability framework

The public purposes of regulation

In section 3 we identified three purposes of regulating online content intermediaries:

- Clarify what consumers should expect from intermediaries, in their handling of harmful and illegal content;
- Ensure intermediaries' governance of online content is proportionate and accountable, and takes a fair and responsible approach to balancing rights;
- In achieving these goals, recognise differences between intermediaries of varying size and different business models, and the need for regulatory certainty and an outcomes-based approach.

In addition, any regulatory model should, as far as possible, be future-proof. It may be impossible to predict what types of content and intermediary will emerge in future, but it is highly likely that policy concerns will continue to arise, as they have since the creation of the Web. New intermediaries will grow and today's platforms may decline.[41]

In this section we describe a model for accountability of online content intermediaries to achieve these goals.

Accountability for outcomes and procedures

Historically, content regulation has tended to the command-and-control end of the regulatory spectrum: detailed rules about what kinds of content are permissible, in what circumstances, and on what services.

This model does not transfer well to the Internet, with its low barriers to entry, increased scope for free expression and vast amounts of content. There are great problems of principle and practice when public agencies seek to decide whether specific content items or providers should be allowed online, as the Chinese example shows.[42]

Instead, as Mark Zuckerberg puts it, intermediaries should be judged by their results.[43] If so, we need means of validating their performance in reducing access to harmful or illegal material, while not inadvertently blocking legal content, and (in certain circumstances) promoting socially preferred content, without unfairly restricting competition.

This approach, known as *outcomes-based accountability*,[44] requires a clear statement of desired outcomes, monitoring of progress towards them, and disclosure of results.

Outcomes-based accountability is likely to require independent evaluation, to ensure credibility and legitimacy. We note that some platforms have signed up to the principle of independent evaluation, for example in the Report of the EU High-Level Group on fake

---

[41] Ofcom has found that apps used by children have changed over time, with Facebook declining, and Snapchat, Instagram, WhatsApp and YouTube rising. Ofcom, *Children and parents: media use and attitudes report 2017*, November 2017. Pew reports similar findings in the US. Pew Research Center, *Teens, Social Media & Technology 2018*, May 2018

[42] CNBC, *China has launched another crackdown on the internet – but it's different this time*, 26 October 2017

[43] Mark Zuckerberg, *oral testimony to a Joint Session of the Senate Judiciary Committee and the Senate Commerce, Science, and Transportation Committee*, April 2018

[44] Mark Friedman, Trying Hard is not Good Enough: How to Produce Measurable Improvements for Customers and Communities, Trafford, 2005

COMMUNICATIONS
CHAMBERS

news, to which Facebook, Google, and Twitter were signatories.[45]

However, outcomes-based accountability may not always be achievable, for example where desired outcomes are impossible to define or measure, or where there are irreconcilable disputes about the appropriate balance between rights.

In such cases *procedural accountability* provides another route to legitimacy, in which intermediaries are judged by whether they used appropriate processes to reach a decision.[46]

Procedural accountability requires a definition of good governance standards and the means to assess whether intermediaries' policies meet those standards.

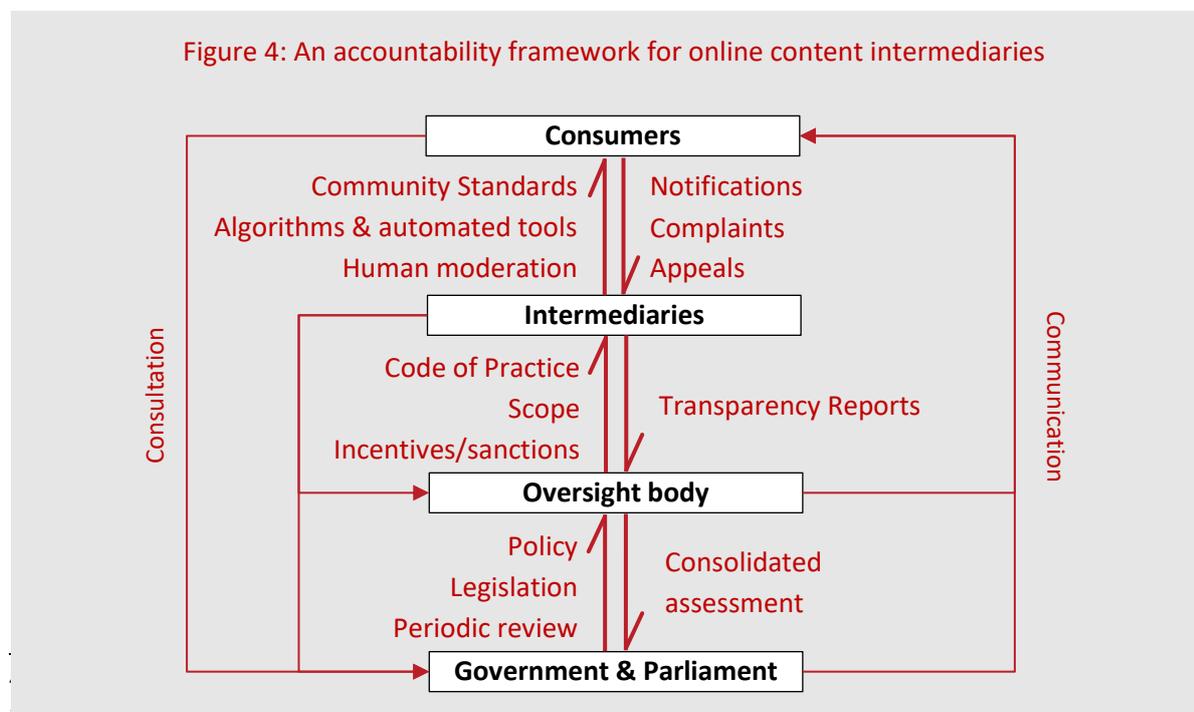## Components of an accountability framework

The proposed accountability framework (summarised in Figure 4) has four core parts:

- A Code of Practice defining overarching content standards and principles of good governance;
- The scope of the Code, and an approach to determining the specific intermediaries that it should apply to;
- An oversight body, tasked with maintaining the Code and the list of intermediaries in scope, requesting certain information, assessing and reporting on compliance with the Code's principles and promoting users' rights;
- Incentives and sanctions, available to the oversight body to encourage compliance.

We discuss each in turn. Two test cases are considered below in Box 5 and Box 6.

Code of Practice

The Government committed to introduce a Code of Practice for social media platforms on online abuse in the Digital Economy Act 2017.[47] It provided a draft Code in the subsequent statement on its Internet Safety

Figure 4: An accountability framework for online content intermediaries

[46] Bunting, *supra* note 1
[47] http://www.legislation.gov.uk/ukpga/2017/30/contents/enacted

Strategy.[48] Importantly, the Code sets standards for both the type of content users should expect platforms to deal with, and the procedures they put in place to do so (including enabling users to report abuse and block abusive users, dealing with user notifications, and communicating their policies and actions).

The scope of the draft Code is limited in three respects: the *type* of content it addresses (for example, unlawful conduct is excluded, despite it having the scope to cause the most harm); the *intermediaries* to which it applies ('social media platforms', although this is not defined in the Act); and the *procedures* it covers (for example it does not address any voluntary use by social media providers of automated blocking or content suppression tools).

Therefore it seems likely that similar Codes may be required in time for other content types, potentially engaging other platforms, and for other kinds of intervention by intermediaries.

It may be considered whether this approach – development of a series of detailed Codes by Government on an ad hoc basis – is sustainable or efficient. Government processes are slow and unwieldy; as things stand, fresh legislation would be required to enable further Codes. As platforms evolve, Codes may need to do so as well, but it is not obvious what the mechanism for this would be.

An alternative approach would be for Government to empower a regulator to provide for a more comprehensive, but higher-level Code of Practice, defining both broad content standards and procedural principles, but with fewer specific requirements.

Content standards

A comprehensive Code would cover a wide range of content types:

- Illegal material, including for example extremist content, hate speech, prohibited images of children, false advertising and intellectual property infringement;
- Legal but harmful material, which should be defined, as precisely as possible, and included in the Code on the basis of an independent materiality assessment showing substantial evidence of harm on intermediary platforms. Possible examples include cyberbullying, misogynistic content, pornography and advertising placed in proximity to unsafe material; and
- Content that meets positive policy goals (such as social inclusion, diversity of news, or provision of 'trustworthy' news[49]).

Given risks of state interference in content, the Code should avoid highly prescriptive rules, especially about content that is not illegal. Intermediaries should have discretion to interpret broad principles as they apply to their particular platforms and users, whose expectations are likely to vary by platform.

The Advertising Standard Authority (ASA)'s Broadcast and Non-Broadcast Codes,[50] the Ofcom Broadcasting Code[51] and ATVOD's Rules & Guidance[52] provide examples of similar codes. Given the diversity of online content, the range of platforms covered, and the need for intermediaries to retain flexibility

---

[48] HM Government, *supra* note 3
[49] Mark Zuckerberg undertook in Jan 2018 to prioritise news on Facebook that was "trustworthy, informative and local"
[50] *https://www.asa.org.uk/codes-and-rulings/advertising-codes.html*
[51] *https://www.ofcom.org.uk/tv-radio-and-on-demand/broadcast-codes/broadcast-code*
[52] *https://www.ofcom.org.uk/__data/assets/pdf_file/0023/39173/a2.pdf*

in line with their particular users' interests, the Code of Practice's content standards might be expected to be relatively less detailed than these documents.

The Code would need to evolve as law evolves, and as fresh evidence emerges about the risks and opportunities of online content.

Procedural standards

Procedural expectations would be based on familiar principles of good governance,[53] and might include:

- Proportionality – intermediaries should only block, remove or suppress content that is demonstrably harmful, illegal, or otherwise contrary to the Code or to their own terms and conditions;
- Evaluation – intermediaries should specify objectives, measure and disclose the impact of their policies and decisions, and make commitments to improve performance over time;
- Transparency – intermediaries' content policies should be prominently available in a user-friendly form, and their impacts disclosed in transparency reports, where appropriate using industry-standard measures of success, which the oversight body could use to publish a consolidated assessment. Disclosure of impacts is important both to provide public confidence and remove information asymmetries between platforms and the oversight body;
- Accessibility – intermediaries should make it easy for users to notify infringing content, give feedback on policies and processes, and access straightforward and quick complaint and appeals processes.

As with content standards, intermediaries could decide what specific procedures to adopt, recognising that these may differ widely between intermediaries.

Scope and obligations

In section 3, we defined intermediaries broadly as services that facilitate exchange between providers and consumers of content and information, including by selecting, sorting, ranking, recommending and suppressing content to users, and that earn revenue by taking a share of value created by exchange.

Social media may be the most common targets of policy attention and public concern. But this broader definition, consistent with European law, encompasses an enormous range of other services, including video aggregators, search engines, streaming sites, bloghosts, discussion forums, online games, publishers with comment functionality, apps stores, messaging services, collaboration tools such as Slack and Github, and more.

An outcomes-based approach entails broad scope: not all online harms begin and end at Facebook, YouTube, and Twitter. But clearly a balance must be struck to avoid requiring a vast number of companies to adopt policies consistent with the Code. Disproportionate regulation and over-reaction to under-cooked theories of harm threaten users' interests in a competitive, innovative and open Internet.

One solution would be to develop a List of intermediaries in scope for the Code, based on evidence of harm and regulatory tiers, analogous to the tiers that exist in broadcast regulation:

---

[53] For example, Better Regulation Taskforce, *Principles of Good Regulation*, 2003

- All firms below a certain de minimis userbase would be excluded from scope of any regulatory action, except where there is evidence of illegal and significantly harmful material on their platforms;[54]
- In Tier 1, intermediaries would be obliged to provide an independent assessment of harm on their platforms, with supporting data, in response to a reasonable, evidence-based and specific request; or to provide data to enable the oversight body to make that assessment. These would be the only Tier 1 obligations;
- In Tier 2, based on a harm assessment, named intermediaries would be required to have a policy that meets the conditions of the Code (or particular parts of it), and to disclose performance against the Code's objectives. A list of intermediaries in scope for Tier 2 would be maintained and made publicly available;
- In Tier 3, in exceptional circumstances and based on clear thresholds for action, requests could be made of specific intermediaries to take particular actions to address substantial harm.

Intermediaries should have a right of appeal if they believe harm assessment requests are disproportionate (Tier 1), the evidence does not support their inclusion in scope (Tier 2), or unnecessary or unsubstantiated obligations have been imposed (Tier 3).

This model seeks to ensure that the scope of regulation is limited only to those intermediaries where there is robust evidence of harm, while also providing flexibility for the companies in scope to change over time and from issue to issue. Expectations would generally be greater for the greatest harms, usually associated with illegal content. Smaller firms should generally be subject to fewer or

---

Box 5. Test case 1: online political advertising and election campaign rules

Online political advertising is not a new phenomenon, although it has only been subject to scrutiny relatively recently. There is no legal basis for the Electoral Commission to require intermediaries to help assess and respond to risks arising from open advertising platforms.

Under our proposal, the Electoral Commission could approach the oversight body for support. If an evidential threshold for investigation were passed, the oversight body could require intermediaries with significant political advertising activity to assess the risk that UK election rules could be breached (for example, by allowing foreign funding of political advertising). Intermediaries would have significant discretion in how they carried out the risk assessment, but the oversight body could specify that it should be independent, and findings would have to be disclosed to the oversight body and the Electoral Commission.

Subject to the risk assessment, the oversight body could require relevant intermediaries to develop a policy to manage the risk. Again, intermediaries would have discretion, but they would need to justify their approach to the oversight body, monitor and disclose its effectiveness over time, and review if it did not meet its intended objectives.

This potential would not have eliminated the risk of foreign interference in UK elections entirely. But it could have resulted in platforms' earlier engagement with the risks, and allowed the development of an appropriately tailored approach.

---

[54] The Government's Response to the Internet Safety Strategy Consultation suggests a Code of Practice to be most relevant to platforms with more than 250,000 UK users

no expectations, in line with their reduced impact and the greater proportionate impact of regulation. As Lorna Woods and William Perrin have pointed out, precedents exist for graduated obligations in other regulatory regimes such as health and safety and data protection.[55]

This model offers flexibility to intermediaries in how they meet the goals of the Code. Many possible models already exist. Intermediaries may act individually, and use whatever combination of automated technology, user notification and human moderation they believe appropriate – as YouTube does today with its ContentID tool for identifying copyright-protected content, or Instagram's efforts to keep its comments troll-free.

Or they might collaborate in industry-led initiatives, potentially including relevant public agencies. The Internet Watch Foundation provides a model, which will not be right for all issues, but may be appropriate when there is benefit from pooling knowledge and a consistent approach.

Looking further ahead, it is possible that a market for 'regulatory services' might develop. Third parties could offering independent measurement, reporting, complaint handling and moderation consistent with the Code, and possibly even with certification from the oversight body. This might provide a cost-effective way for smaller intermediaries to deal with any particular issues identified on their platforms.

Oversight

Achieving regulatory certainty within a flexible regime requires ongoing oversight and iteration of the Code. Users and intermediaries will also need advice on their rights and responsibilities. An oversight body will meet these needs, while operating as efficiently as possible and without unnecessary bureaucracy.

Oversight functions should be specific and limited. They could include:

- Developing, maintaining and publicising the Code, and communicating to platform users their rights and responsibilities, and what they should expect of intermediaries;
- Maintaining the List of intermediaries in scope for some or all of the Code, and requiring or carrying out harm assessments;
- Applying regulatory incentives and sanctions where necessary to secure compliance with the Code (see below);
- Providing a consolidated assessment of the effectiveness of the Code in meeting its statutory aims;
- Potentially, acting as a backstop for complaints that cannot be resolved by intermediaries themselves, although only if this can be done without unduly undermining intermediaries' autonomy or overloading the oversight body with enormous numbers of complaints.

The oversight body and the Code itself would be independent of both Government and industry. Government would not be able unilaterally to add particular content types to the Code or intermediaries to the List.

Instead, the oversight body should be tasked with consulting with relevant Government departments and agencies in developing the Code. The range of issues covered may be broad, ranging from (for example) hate speech and extremism, through bullying, intimidation

---

[55] Lorna Woods and William Perrin, *Harm Reduction In Social Media – What Can We Learn From Other Models Of Regulation?*, Carnegie Trust, May 2018

and domestic violence, to accuracy in advertising and copyright. A consistent approach is needed.

The oversight body would also be responsible for consulting on the Code and any other interventions, and for providing a means of appeal for intermediaries who believe they have been wrongly included in scope.

Any online content intermediary with more than the de minimis number of UK users could be required to notify the oversight body, as in the notifications regimes maintained by Ofcom for providers of electronic communications networks or services,[56] and on-demand programme service providers.[57]

We discuss institutional options for the oversight body in section 6.

Incentives and sanctions

This regime stands or falls on its ability to secure intermediaries' engagement with and support for the Code of Practice. This in turn depends on appropriate and proportionate incentives and sanctions.

The oversight body might offer incentives to encourage compliance: accreditation, kitemarks, beneficial rights (e.g. to access adjudication or arbitration mechanisms). These incentives should be developed with input from intermediaries.

Should the incentives be insufficient, the oversight body should also have sanctions available to it, potentially including the ability to: issue warnings; impose fines; provide notices to third parties who provide services to intermediaries (e.g. payment providers or advertisers); and, in extreme cases, involving repeated failures to comply, the power to request ISPs to block services.[58]

The most onerous sanctions should be restricted to cases where intermediaries have a strong interest in providing access to illegal

---

Box 6. Test case 2: online bullying

In this case, there is no legal definition of harmful activity, and potential cost to online freedoms if the definition is drawn too broadly. Consequently careful definition and analysis of the nature and extent of alleged problems is needed.

The oversight body could require intermediaries to assess the risk and impact of bullying on their platforms, including reviewing and if necessary commissioning independent expert evidence. The oversight body could play a role in creating advisory forums enabling companies and external stakeholders to discuss possible approaches, at a pan-industry level where appropriate.

Any platforms on which bullying was identified as a significant problem could be required to develop and publish anti-bullying policies, specify measures of success, and disclose the effectiveness of their actions including the efficacy of any takedown procedures. The oversight body would not be empowered to dictate specific policies, but could require that they were independently validated, and might have a statutory duty to report on their effectiveness, on a comparable cross-industry basis by default.

---

[56] Communications Act 2003, *section 33*
[57] Communications Act 2003, *section 368BA*
[58] As in the Digital Economy Act provisions in relation to age verification by providers of online pornography. Digital Economy Act 2017, *section 21*

or harmful material, have the capacity to prevent it, and repeatedly decline to take proportionate action.

As Woods and Perrin highlight, the power to block a service poses risks to free speech and could have significant consumer impact. Any such powers would need to be carefully circumscribed.[59]

Interaction with the liability regime established under the E-Commerce Directive, and any potential future liability or duty of care arrangements, would also need careful consideration. These issues are discussed in section 7.

Conclusion: the value of accountability

The premise of the approach described here is simple: requiring intermediaries to assess risks, and take action where they are identified, will over time lead to reduction of those risks, without constraining platforms' independence and autonomy. Indeed it is possible that intermediaries would take appropriate steps *prior* to the involvement of the oversight body, to seek to avoid being caught in a regulatory process.

It will be impossible to eliminate risks entirely, but this approach encourages intermediaries to engage earlier with potential problems, internalise social impacts that they may not have foreseen or understood, and reach an appropriate balance between the various interests in play, including their and their users' commercial considerations.

---

[59] Woods & Perrin, *supra* note 55

6. Legislation and options for the oversight body

Self-regulation vs co-regulation

Self-regulation, co-regulation and statutory regulation are often seen as three distinct models, but the lines are blurry, and in practice most regulatory regimes have some mix of statutory underpinning and industry discretion. In this case, intermediaries will need to be closely involved in the development and operation of the accountability framework. The question is what mix of statutory and intermediary responsibility is most appropriate.

An enabling framework in statute

Forthcoming legislation could provide the foundation for the framework described here. Legislation would be needed to:

- Make provision for an oversight body;
- Specify its responsibilities and powers, including to develop the Code of Practice, assess online harms (including by requiring intermediaries to provide information and analysis), and require intermediaries to develop policies to mitigate harm. Legislation would also define thresholds for different tiers of intervention by the oversight body;
- Describe in broad terms what the Code should cover;
- Define in law the concept of an 'online content intermediary'; and
- Provide a statutory basis for incentives and sanctions.

Options for the oversight body

Within this statutory framework, there are three broad options for the oversight body, each with pros and cons.

First, **an industry co-regulator**. The creation of an oversight body may benefit intermediaries,

as long as it operates within clearly defined constraints and in an accountable, proportionate way. It may insulate them against the potential for more intrusive, ad hoc or costly regulation. It may support their efforts to rebuild reputational capital after the damaging controversies of recent years, helping with customer and staff retention. It may be particularly well-placed to develop effective consumer communications, improving the public's understanding of intermediaries policies, and thereby building trust.

Therefore, Government might consider whether legislation should allow for the industry to create, fund and operate an independent co-regulatory body to provide the oversight functions defined in statute. And the industry should consider whether it is now the right time to develop such a proposal.

The Advertising Standards Authority (ASA) illustrates the potential for industry-led bodies to secure consumer, government and stakeholder trust. As with the ASA, any intermediary industry body would need to maintain robust standards of independence, accountability and consumer focus, and have binding powers of enforcement.

A public body would need to have backstop powers, as Ofcom does for the ASA's Broadcast Code of Practice, for example to sign off major changes to the Code of Practice. It may also be appropriate for legislation to require periodic independent reviews to enable Government and Parliament to assess its effectiveness.

This option has significant benefits in terms of independence, flexibility and expertise, but it would have to run at arms-length from the industry (for example with industry

representation via an advisory group, not on the main board), and in line with public purposes defined in statute.

This option would be compromised if the industry body were seen as insufficiently independent, dominated by bigger firms or unable to secure participation of all relevant intermediaries.

The second option is for an **existing statutory body** to provide oversight functions. Woods and Perrin consider the options and conclude Ofcom is the best placed of the existing regulators.[60] It has relevant experience, familiarity with content regulation, a strong research capability, a track record of independence, a consumer panel and resilience in dealing with big firms.

These arguments are strong, but set against them we note that Ofcom already has a very extensive brief, and is primarily a sectoral regulator for the telecoms and media sectors. The oversight body would be expected to cover a wide range of issues including many with which Ofcom has little expertise, but where other public bodies do (for example, harassment, extremism, advertising, copyright). It would have to work in new ways with a much wider range of stakeholders.

The third option is to create a **new statutory body**. Woods and Perrin argue this would take too long[61] (although it seems to us that the main barrier to rapid implementation is the likely speed of the legislative process, not the creation of the regulator). However, this might be the most costly option, and there would be concerns about whether a new public body would have the scale, competence and legitimacy to engage with global

intermediaries on equal terms. One possibility may be for the proposed Centre for Data Ethics and Innovation to take on the oversight task, depending on final decisions about its appropriate role and responsibilities, on which the Government is consulting at the time of writing.[62]

In either of the latter two options, the industry could still be expected to fund the statutory body, via fees paid by notifying intermediaries, similar to the way Ofcom is funded by its licensees. Woods and Perrin suggest, alternatively, that the costs could be met from a share of revenue planned to be raised by the Government from taxing internet company revenues, should this proposal be implemented.

Conclusion: conditions for successful intermediary oversight

Oversight provides clarity about the expectations of platforms, both for consumers and intermediaries themselves. Given the diversity and pace of change of online markets, flexibility is essential.

Oversight has risks. The oversight body may impose requirements which are not supported by evidence of real harm. Its rules may have impacts on openness and innovation in content markets, constrain freedom of speech, and lock in the power of big platforms. The oversight body may struggle to acquire legitimacy, whether it is a statutory body or established by the industry. It may be too influenced by political pressure or lobbying.

A statutory body may be particularly vulnerable to accusations that it is too remote from platform activities or users' needs; an

---

[60] Lorna Woods and William Perrin, *Who Should Regulate To Reduce Harm In Social Media Services?*, Carnegie Trust, May 2018

[61] *Ibid.*

[62] Department for Digital, Culture, Media and Sport, *Consultation on the Centre for Data Ethics and Innovation*, June 2018

industry body may be criticised for being insufficiently independent.

To some extent these risks can be mitigated by establishing guardrails in law around the oversight body's duties and powers. Legislation should require periodic reviews of the oversight body, ideally independent of Government, and for mechanisms of Parliamentary accountability.

But in the end the legitimacy of intermediary oversight, and possibly the activities of platforms themselves, will depend on how good the oversight is. It is important to be clear that the kind of oversight described here will not solve all content problems online. The issues are hard, requiring an iterative, collaborative approach. It is impossible, without very significant unintended consequences, to iron out all occurrences of illegal and harmful content. Intermediaries should be in the front line of efforts to reduce exposure to unwanted content, not regulators.

But these considerations make accountability more important, not less. At present, online content regulation is largely the preserve of private intermediaries. The intent of these proposals is to provide more systematic scrutiny and accountability of that activity, recognising that perfection is impossible, but that consumers and citizens have a right to know how content is being managed in their interests.

[blank page]

7. Conclusions and recommendations

Online content regulation is not impossible. The question is how

No country has yet put in place a systematic solution to intermediary accountability for online content, that helps users understand what they can expect across a wide range of types of content and platform, and ensures that intermediary actions are proportionate and linked to evidence of harm.

This paper has tried to show such a solution is possible. The question, as Mark Zuckerberg says, is not whether to regulate, but how.[63]

The framework we have described draws on existing law, and policy proposals already under development here and elsewhere.

The Government has an opportunity, in the White Paper on online harms and safety, to establish this framework. Our hope is that this paper contributes to the ongoing policy debate. Much work is needed, of course, with both platforms and stakeholders, on detailed questions of definition, scope, thresholds for intervention and the oversight role.

Considerations for implementation

Two particular issues need consideration in implementing an accountability framework.

First, **consistency with the ECD**, in particular Article 14, which establishes liability for content hosts in certain limited conditions, and Article 15, which prevents member states from imposing 'general monitoring obligations' on intermediaries.

We believe our proposal is consistent with these provisions. We have not proposed that liability should arise in different conditions nor

that the existing limitations should change. Nor could a regulator introduce general monitoring obligations under our proposal; it would be for intermediaries to develop policies and governance practices consistent with the overall goals and principles of the Code of Practice, which may include general monitoring but it is not required. For the avoidance of doubt, legislation could prohibit any such monitoring requirements.

The ECD is a foundation, but it explicitly does not preclude other legal and regulatory activity. For example Recital 48 anticipates member states may apply duties of care on intermediaries to detect and prevent certain types of illegal activities. The ECD also requires member states to encourage industry associations or organisations to draw up codes of conduct to contribute to the implementation of the Directive (Article 16). In a number of areas, as described in section 4, European policy is itself moving towards more proactive obligations on intermediaries.

The Commission has said that intermediaries which take proactive steps to prevent hate speech and other harmful content should not be regarded as assuming liability for it.[64]

Longer-term, it may be appropriate to review the ECD's liability conditions. These conditions may discourage intermediaries for introducing the most effective policies in case they inadvertently open them to liability under the 'actual knowledge' test. A review might consider the possibility of linking liability shields instead to a wider notion of platform accountability, that requires intermediaries to assess the impact of their policies and mitigate harms across a range of issues, as we have proposed for the Code of Practice.

---

63 Mark Zuckerberg, _oral testimony to House of Representatives Energy and Commerce Committee_, 11 April 2018

64 European Commission, _Communication on Tackling Illegal Content Online_, 28 September 2017

COMMUNICATIONS
CHAMBERS

Similarly, should a duty of care be introduced for content intermediaries, it may be appropriate to link the required standard of care to compliance with the Code.

This approach could be rooted in European human rights law. Member states have a positive obligation to protect human rights, including protecting individuals from private parties' actions by ensuring compliance with relevant legislative and regulatory frameworks. Given the growing importance of the Internet to the fulfilment of rights, and the role of intermediaries in balancing rights, there is a strong case for putting in place oversight mechanisms to ensure they provide appropriate protections.

Second, **scope for unilateral action** by the UK. It is often argued that individual countries cannot regulate global platforms. There are risks that unilateral action creates a 'splinternet'. But while it may be preferable to oversee intermediaries across multiple jurisdictions, that does not prevent individual countries from acting – as Germany already has. Internet companies already provision some content on a country by country basis, making local regulation practical.

In leaving much discretion to intermediaries, our proposal reduces the risk of conflict with other jurisdictions. Many of the content and procedural standards that we have suggested for the Code already exist in domestic and European law. The costs of national variation in content policies can be treated as a legitimate factor in assessing intermediaries' accountability. The UK would not need to enforce its standards on other countries, as the German law may do.[65]

One advantage of the industry organising to form an independent oversight body, as suggested in section 6, is that this body could work internationally and provide a forum for intermediaries to engage with multiple governments. The Global Internet Forum to Counter Terrorism provides a precedent of a sort, although it does not have any statutory or co-regulatory status.[66]

Recommendations

Government should include an accountability framework for online content intermediaries in the planned White Paper on online harms and safety. This should provide for:

- a Code of Practice that describes desired standards across a range of content types, and procedural expectations of platforms, including Transparency Reports; and
- An oversight body responsible for developing the Code and with the means to deploy incentives and sanctions to encourage the Code's take-up.

Second, intermediaries should work together to assess the potential for a co-regulatory body to provide independent oversight of intermediaries' content policies, with buy-in from most platforms with significant numbers of UK users. It would be sensible for industry to consider whether such a body could operate across multiple jurisdictions, and consequently help address regulatory concerns in other countries as well as the UK.

Finally, Government should consider options for a statutory oversight body, in case the industry option does not make sufficient progress within a reasonable time period.

---

[65] A recent German court case determined that geoblocking illegal content based on IP addresses was insufficient to meet NetzDG's requirement to prevent its availability to users in Germany. Platforms may respond by blocking content that contravenes German law everywhere, even though Germany has significantly stronger speech laws than most countries. The Atlantic, _Germany's Attempt to Fix Facebook Is Backfiring_, 18 May 2018

[66] _https://www.blog.google/topics/google-europe/update-global-internet-forum-counter-terrorism/_