HM Government

Online Harms White Paper: Full Government Response to the consultation

December 2020

CP 354

Online Harms White Paper: Full Government Response to the consultation

Presented to Parliament
by the Secretary of State for Digital, Culture, Media & Sport
and the Secretary of State for the Home Department
by Command of Her Majesty

December 2020

CP 354

This publication is available at www.gov.uk/official-documents.

Any enquiries regarding this publication should be sent to us at ohwp.fullgovernmentresponse@dcms.gov.uk

ISBN 978-1-5286-2312-4

CCS1220695430 12/20

Printed on paper containing 75% recycled fibre content minimum

Printed in the UK by the APS Group on behalf of the Controller of Her
Majesty’s Stationery Office

Table of contents

Joint Ministerial foreword

Our world is now a digital one. From connecting with loved ones, to the way we do business and deliver public services - almost every part of our lives is at least now partly online.

But the COVID-19 pandemic has shone a spotlight on the risks posed by harmful activity and content online. The pandemic drove a spike in disinformation and misinformation, and some people took advantage of the uncertainty to incite fear and cause confusion. The pandemic has also underlined a much more grave problem; the risks posed to children online. In a month-long period during lockdown, the Internet Watch Foundation and its partners blocked at least 8.8 million attempts by UK internet users to access videos and images of children suffering sexual abuse.[1]

This government is unashamedly pro-tech. We are committed to using digital technologies and services to power economic growth across the entire UK, and ensuring a more inclusive, competitive and innovative digital economy for the future. We are taking action to unlock innovation across digital markets, while also ensuring we keep people safe online and promote a thriving democracy, where pluralism and freedom of expression are protected. To unleash growth we need to ensure there is trust in technology.

The government's response to online harms is a key part of our plans to usher in a new age of accountability for tech companies, which is commensurate with the role they play in our daily lives. Our ambition is to build public trust in the technologies that so many of us rely on. Ultimately, we must be able to look parents in the eye and assure them we are doing everything we can to protect their children from harm.

This response to the Online Harms White Paper sets out plans for a new duty of care to make companies take responsibility for the safety of their users. It builds on our manifesto commitment to introduce legislation to make the UK the safest place in the world to be online but at the same time defend freedom of expression.

The legislation will define what harmful content will be in scope. Principally, this legislation will tackle illegal activity taking place online and prevent children from being exposed to inappropriate material. But the legislation will also address other types of harm that spread

[1] '[Millions of attempts to access child sexual abuse online during lockdown](#)' Internet Watch Foundation, 2020 (last viewed in November 2020)

online - from dangerous misinformation spreading lies about vaccines to destructive pro-anorexia content.

These new laws will mean no more empty gestures - we will set out categories of harm in secondary legislation and hold tech giants to account for the way they address this content on their platforms. This approach will empower people to manage their online safety and ensure that these companies will not be able to arbitrarily remove controversial viewpoints.

Alongside tackling harmful content this legislation will protect freedom of expression and uphold media freedom. Companies will be required to have accessible and effective complaints mechanisms so that users can object if they feel their content has been removed unfairly.

And this regulation will be proportionate. Fewer than 3% of UK businesses will be in scope. We will focus on the biggest, highest risk online companies where most illegal and harmful activity is taking place.

This groundbreaking regulatory framework will be enshrined in law through the upcoming Online Safety Bill.

Our criminal law must also be fit for the digital age and provide the protections that victims deserve. The Law Commission is currently reviewing whether new offences are necessary to deal with emerging issues such as cyber-flashing and 'pile-on' harassment. We will carefully consider using the online harms legislation to bring the Law Commission's final recommendations into law, where it is necessary and appropriate to do so.

As an independent country, the UK has the opportunity to set the global standard for a risk-based, proportionate regulatory framework that protects citizens online and upholds their right to freedom of expression. We will work with our international partners to develop common approaches to this shared challenge, whilst delivering on our ambition to make the UK the safest place in the world to go online. We will lead, but we are confident others will join us.

Rt Hon Oliver Dowden CBE MP	Rt Hon Priti Patel MP
Secretary of State for Digital, Culture, Media and Sport	**Secretary of State for the Home Department**

Executive summary

1. The Online Harms White Paper set out the government's ambition to make the UK the safest place in the world to go online, and the best place to grow and start a digital business. It described a new regulatory framework establishing a duty of care on companies to improve the safety of their users online, overseen and enforced by an independent regulator. This will build public trust in the services that these companies are offering, and support a thriving and fast-growing digital sector. The White Paper proposed that regulation be proportionate and risk-based, ensuring companies have appropriate systems and processes in place to tackle harmful content and activity. It also made clear that the framework will protect users' rights, including freedom of expression online.

2. The government set out the results of the formal consultation and clarified its direction of travel in the Online Harms White Paper - Initial government response, published in February 2020. The initial government response reconfirmed our commitment to the duty of care approach set out in the White Paper and announced a number of further measures to increase proportionality and protect freedom of expression. It also indicated that the government was minded to appoint Ofcom as the regulator. The government has continued to develop its policy proposals since February and has made further, important changes. The full government response confirms that Ofcom will be named as the regulator in legislation, and sets out the intended policy position.

3. The government has taken a deliberately consultative and iterative approach in developing the framework, to ensure regulation that is coherent, proportionate and agile in response to advances in technology. It is part of the government's overarching, pro-innovation approach to regulating digital technologies, that will address issues arising from digital technology which affect prosperity, security and our democratic values. This is an important step forward in building a safer and more prosperous digital future for everyone.

4. Tackling online harms is a global problem and the government recognises that legislation and regulation in the UK, and elsewhere, forms only part of the response required. The UK, with its strengths in digital innovation, highly respected legal system, business-friendly environment and world-class regulators, has an opportunity to act as a global leader in this space. That is why the government is working closely with many of our international partners to address this shared challenge in order to work towards common approaches to tackling online harms. The development of the online harms regime represents an important step in the UK's strategy to create a coherent and pro-innovation framework for the governance of digital technologies, and to set the global standard for a risk-based, proportionate regulatory framework.

The continuing case for action

5. The internet has, in many ways, transformed our lives for the better. It has revolutionised our ability to connect with each other and created previously inconceivable economic opportunities. Internet use in the UK across all adult age groups increased from 80.9% in

2012 to 90.8% in 2019.[2] In April 2020, internet users in the UK spent an average of 4 hours 2 minutes online each day, a record figure.[3]

6. However, the case for robust regulatory action continues to grow. Over three quarters of UK adults express a concern about going online,[4] and fewer parents feel the benefits outweigh the risks of their children being online, with the proportion falling from 65% in 2015 to 55% in 2019.[5]

7. The White Paper set out the extensive evidence of illegal and harmful content and activity taking place online. The government highlighted the prevalence of the most serious illegal harms which threaten our national security and the physical safety of children. It also explained how online services are being used as a tool for abuse. The White Paper acknowledged growing concerns about the impact of harmful content on the wellbeing of children in particular. These problems have not gone away.

8. In terms of illegal content and activity, there were more than 69 million images and videos related to child sexual exploitation and abuse referred by US technology companies to the National Center for Missing and Exploited Children in 2019,[6] an increase of more than 50% on the previous year.[7] In 2019, of the over 260,000 reports assessed by the Internet Watch Foundation, 132,730 contained images and/or videos of children being sexually abused (compared to 105,047 in 2018), and 46% of reports involved imagery depicting children who appeared to be 10 years old or younger.[8] Between its launch in January 2015 and March 2019, 8.3 million images have been added to the Child Abuse Image Database.[9] The National Crime Agency estimates at least 300,000 individuals in the UK pose a sexual threat to children.[10]

9. Terrorist groups use the internet to spread propaganda designed to radicalise, recruit and inspire vulnerable people, and to incite, provide information to enable, and celebrate terrorist attacks. Some companies are taking positive steps to combat online terrorist content. The larger platforms are already taking proactive measures and using automated technology. For instance, Twitter actioned 95,887 unique accounts related to the promotion of terrorism/violent extremism between January and June 2019.[11] However, terrorists and

[2] 'Internet users' Office for National Statistics, May 2019 (last viewed in November 2020) "Internet use" here refers to respondents who have used the internet in the last three months"

[3] 'Online Nation: narrative report' Ofcom, June 2020 (last viewed in November 2020)

[4] 'Internet users' concerns about and experience of potential online harms' Ofcom and ICO, May 2019 (last viewed in November 2020)

[5] 'Children and parents: Media use and attitudes report 2019' Ofcom, February 2020 (last viewed in November 2020)

[6] 'CyberTipline' National Center for Missing and Exploited Children (last viewed in November 2020)

[7] 'Tech Companies Detect a Surge in Online Videos of Child Sexual Abuse' The New York Times, February 2020 (last viewed in November 2020)

[8] 'The Internet Watch Foundation Annual Report 2019' The Internet Watch Foundation, April 2020 (last viewed in November 2020)

[9] 'Child sexual abuse - Appendix tables' Office for National Statistics, January 2020 (last viewed in November 2020)

[10] 'Law enforcement in coronavirus online safety push as National Crime Agency reveals 300,000 in UK pose sexual threat to children' National Crime Agency, April 2020 (last viewed in November 2020)

[11] 'Rules Enforcement' Twitter, August 2020 (last viewed in November 2020)

their supporters continue to use a wide range of platforms to further their aims. It is critical that industry works together, and that the government and industry continue to build on the foundations laid by initiatives such as the Global Internet Forum to Counter Terrorism, to prevent exploitation of the internet for terrorist purposes.

10. Alongside illegal content and activity, the White Paper highlighted increasing levels of public concern about online content and activity which is lawful but potentially harmful. This type of activity can range from online bullying and abuse, to advocacy of self-harm, to spreading disinformation and misinformation. Whilst this behaviour may fall short of amounting to a criminal offence it can have corrosive and damaging effects, creating toxic online environments and negatively impacting users' ability to express themselves online.

11. In 2019, according to research conducted by Ofcom and the Information Commissioner's Office, 23% of 12-15 year olds had experienced or seen bullying, abusive behaviour or threats on the internet in the last 12 months.[12] Nearly half of girls admit to holding back their opinion on social media for fear of being criticised.[13] Galop, the LGBT+ anti-violence charity's, most recent online hate crime survey highlighted that 8 in 10 respondents had experienced anti-LGBT+ online abuse in the last 5 years.[14] In 2019, the Community Security Trust, a charity that protects British Jews from antisemitism, saw a 50% rise in reported anti-Semitic online incidents compared to 2018.[15]

12. During the COVID-19 pandemic, digital technologies have brought huge benefits - from unlocking innovation across public services, to enabling millions to work remotely, to supporting people to stay in touch with their friends and families. However, the risks posed by illegal and harmful content and activity online have also been thrown into sharp relief as digital services have played an increasingly central role in our lives.

13. Research shows that 47% of children and teens have seen content that they wished they hadn't seen during lockdown.[16] In a month-long period during lockdown, the Internet Watch Foundation and its partners blocked at least 8.8 million attempts by UK internet users to access videos and images of children suffering sexual abuse.[17] The pandemic also drove a spike in disinformation (the deliberate creation and dissemination of false and/or manipulated information that is intended to deceive and mislead audiences) and misinformation (inadvertently sharing false information) online. Social media has been the biggest source of false or misleading information about 5G technologies and COVID-19 vaccinations during the pandemic.[18]

[12] 'Internet users' concerns about and experience of potential online harms' Ofcom and ICO, May 2019 (last viewed in November 2020)

[13] 'Reclaiming the Internet for Girls' Plan International (last viewed in November 2020)

[14] 'Online Hate Crime Report 2020' Galop (last viewed in November 2020)

[15] 'Antisemitic Incidents Report 2019' Community Security Trust (last viewed in November 2020)

[16] 'Half of children and teens exposed to harmful online content while in lockdown' BBFC, May 2020 (last viewed in November 2020)

[17] 'Millions of attempts to access child sexual abuse online during lockdown' Internet Watch Foundation, 2020 (last viewed in November 2020)

[18] 'Covid-19 news and information: consumption and attitudes - interactive data' Ofcom, June 2020: Week 10-25 of survey Q10c and Q10e (last viewed in November 2020)

14. Many of the major social media companies have moved further and faster than ever before to tackle disinformation and misinformation during the pandemic through technical changes to their products, including techniques to protect user safety online**.** However, this is inconsistent across services. The new regulatory framework will create incentives to ensure that companies continue to take consistent and transparent action to keep their users safe. COVID-19 has shone a spotlight on the need to better understand and respond to new and evolving challenges online, particularly the risks posed to children.

Our response

15. The government's approach to the governance of digital technologies aims to maximise the benefits while minimising the risks. Action is being taken in a range of areas - including data and data use, cyber security, competition, and protecting quality journalistic content - to improve online safety and security, support dynamic and competitive digital markets, and to promote our democratic values online. Our approach is proportionate with innovation at its heart. A future digital strategy will set out how the government is bringing these strands of work together.

16. The government's response to online harms is a key part of this overall approach. The online harms regime will improve users' safety online, build public trust in digital services, support innovation and drive digital and economic growth.

17. The online harms framework will be coherent and comprehensive, bringing much needed clarity to the regulatory landscape and providing support for both industry and users. It will be proportionate, risk-based and tightly defined in its scope. The legislation will avoid taking a 'one size fits all approach' to companies and harms in scope, to reflect the diversity of online services and harms. The government has placed particular emphasis on protecting children,[19] ensuring a pro-innovation approach, and protecting freedom of expression online. Regulation will safeguard pluralism and ensure internet users can continue to engage in robust debate online.

18. Regulation will be only one part of the solution. The government will support growth and innovation across the UK's safety tech sector, creating the right conditions for UK safety tech companies to deliver cutting edge safety technologies. Users must also be empowered to think critically about what they encounter online, and online products and services must be designed from the outset to be safe for users.

Overview of the new regulatory framework for online harms

Which online services will be in scope of the new regulatory framework?

Services in scope and exemptions

19. The new regulatory framework will apply to companies whose services:
    (a) host user-generated content which can be accessed by users in the UK; and/or

[19] For the purposes of this document, in the context of online harms legislation, 'children' means individuals under 18.

(b) facilitate public or private online interaction between service users, one or more of whom is in the UK.

It will also apply to search engines.

20. The legislation will apply to any in-scope company that provides services to UK users, regardless of where it is based in the world. Only a small proportion of UK businesses (the government estimates fewer than 3%)[20] will fall within the scope of the legislation following the new exemptions set out below. Ofcom's regulatory approach will focus on companies where the risk of harm is greatest.

21. The initial government response confirmed that business-to-business services would be out of scope. Services which play a functional role in enabling online activity, such as internet service providers, will also be exempt from the duty of care, although they will have duties to cooperate with the regulator on business disruption measures. The government is introducing additional provisions to exempt many low-risk businesses from the duty of care altogether. New exemptions include services used internally by businesses, and many low-risk businesses with limited functionality (for example retailers who offer only product and service reviews). This avoids imposing regulatory burdens on low-risk companies.

Journalistic content

22. Stakeholders raised concerns during the consultation about how the legislation will impact journalistic content online and the importance of upholding media freedom. Content published by a news publisher on its own site (e.g. on a newspaper or broadcaster's website) will not be in scope of the regulatory framework and user comments on that content will be exempted.

23. In order to protect media freedom, legislation will include robust protections for journalistic content shared on in-scope services. The government is committed to defending the invaluable role of a free media and is clear that online safety measures must do this. The government will continue to engage with a range of stakeholders to develop our proposals.

What harmful content or activity will the new regulatory framework apply to, and what action will companies need to take?

Definition of harm

24. The legislation will set out a general definition of harmful content and activity. A limited number of priority categories of harmful content, posing the greatest risk to users, will be set out in secondary legislation. This will provide legal certainty for companies and users.

Duty of care and the principles of the regulatory framework

25. Under the new legislative framework, companies in scope will have a duty of care towards their users. The legislation will require companies to prevent the proliferation of illegal content and activity online, and ensure that children who use their services are not exposed

[20] DCMS Online Harms research (externally commissioned), 2020, publication date tbc.

to harmful content. It will also hold the largest tech companies to account for what they say they are doing to tackle activity and content that is harmful to adults using their services. Further details on the approach are set out in paragraphs 27 and 28 below.

26. To meet the duty of care, companies in scope will need to understand the risk of harm to individuals on their services and put in place appropriate systems and processes to improve user safety. Ofcom will oversee and enforce companies' compliance with the duty of care. Companies and the regulator will need to act in line with a set of guiding principles. These include improving user safety, protecting children and ensuring proportionality. Further details are set out in Annex A.

Differentiated expectations on companies

27. The regulatory framework will establish differentiated expectations on companies in scope with regard to different categories of content and activity on their services: that which is illegal; that which is harmful to children; and that which is legal when accessed by adults but which may be harmful to them.

28. The new regulatory framework will take a tiered approach. The vast majority of services will be 'Category 2 services'. These companies will need to take proportionate steps to address relevant illegal content and activity,[21] and to protect children. A small group of high-risk, high-reach services will be designated as 'Category 1 services', and only providers of these services will additionally be required to take action in respect of content or activity on their services which is legal but harmful to adults. This tiered approach will protect freedom of expression and mitigate the risk of disproportionate burdens on small businesses. It will also ensure that companies with the largest online presence are held to account, addressing the mismatch between companies' stated safety policies and many users' experiences online.

Public and private communications channels

29. The regulatory framework will apply to public communication channels and services where users expect a greater degree of privacy - for example online instant messaging services and closed social media groups. Ofcom will set out how companies can fulfil their duty of care in codes of practice, including what measures are likely to be appropriate in the context of private communications. This could include steps to make services safer by design, such as limiting the ability for anonymous adults to contact children. Companies in scope will need to consider the impact on users' privacy and ensure users understand how company systems and processes affect user privacy.

30. The scale, severity and complexity of child sexual exploitation and abuse is particularly concerning, with private channels being exploited by offenders. For example, 12 million of the 18.4 million worldwide child sexual exploitation and abuse reports made by Facebook

[21] For ease we have referred to illegal content and activity that meets the definition of harm (see paragraph 2.24) as 'relevant illegal content and activity'.

in 2019 were for content shared on private channels.[22] In light of this, the regulator will have the power to require companies to use automated technology that is highly accurate to identify illegal child sexual exploitation and abuse content or activity on their services, including, where proportionate, on private channels. Recognising the importance of users' privacy, the government will ensure this will be subject to stringent legal safeguards to protect users' rights. The regulator will advise the government on the accuracy of tools and make operational decisions regarding whether or not a specific company should be required to use them. However, before the regulator can use these powers it will need to seek approval from Ministers on the basis that sufficiently accurate tools exist. The regulator will also be able to require companies to use highly accurate technology to identify illegal terrorist content, also subject to stringent safeguards but on public channels only.

Codes of practice

31. Ofcom will issue codes of practice which outline the systems and processes that companies need to adopt to fulfil their duty of care. Companies will need to comply with the codes, or be able to demonstrate to the regulator that an alternative approach is equally effective. The government will set objectives for the codes in legislation. Ofcom will have a duty to consult on the codes, and must help all companies to understand and fulfil their responsibilities. Ofcom must also publish an economic impact assessment for each code and will have a specific duty to assess the impact of its proposals on small and micro businesses, to avoid undue regulatory burdens.

32. The government is publishing interim codes on terrorism and child sexual exploitation and abuse alongside this response, due to the seriousness of these illegal harms. These voluntary and non-binding interim codes will help companies begin to implement the necessary changes and bridge the gap until Ofcom issues its statutory codes of practice.

Additional duties on companies

33. All companies in scope will have a number of additional duties beyond the core duty of care. These include providing mechanisms to allow users to report harmful content or activity and to appeal the takedown of their content. All companies providing Category 1 services will be required to publish transparency reports containing information about the steps they are taking to tackle online harms on those services. The Secretary of State for the Department of Digital, Culture, Media and Sport will have the power to extend the scope of companies who will be required to publish transparency reports, beyond Category 1 companies, if necessary.

Disinformation and misinformation

34. Disinformation and misinformation that could cause significant harm to an individual will be within scope of the duty of care. Some types of disinformation and misinformation are likely to be proposed in secondary legislation as categories of priority harm that companies must

[22] 'Tech Companies Detect a Surge in Online Videos of Child Sexual Abuse' The New York Times, February 2020 (last viewed in November 2020)

address in their terms and conditions. In addition to the requirements under the duty of care, the legislation will introduce further provisions to address the evolving threat of disinformation and misinformation. This will include specific transparency requirements and the establishment of an expert working group, targeted at building understanding and driving action to tackle these issues.

How will the independent regulator oversee and enforce the new regulatory framework?

The regulator

35. Ofcom will be named as the independent regulator in the legislation. Ofcom is a well-established independent regulator with a strong reputation internationally and deep experience of balancing prevention of harm with freedom of speech considerations. It has a proven track record of taking evidence-based decisions, which balance robust consumer protection with the need to ensure the regulatory environment is conducive to economic growth and innovation. This makes it a strong strategic fit for the role.

36. Ofcom will cover the costs of running the regime from industry fees. Only companies above a threshold based on global annual revenue will be required to notify and pay the fees. In practice, this means that a large proportion of in-scope companies will be exempt from paying a fee.

Functions of the regulator

37. Ofcom will have a range of duties and functions under the framework. Its primary duty will be to improve the safety of users of online services (and that of non-users who may be directly affected by others' use of them). This will include setting codes of practice, establishing a transparency, trust and accountability framework and requiring all in-scope companies to have effective and accessible mechanisms for users to report concerns. Ofcom will also have a legal duty to pay due regard to innovation, which will be underpinned by a number of non-legislative measures.

38. To ensure the effective implementation of the regime, Ofcom will have robust enforcement tools to tackle non-compliance, including the power to issue fines of up to £18 million or 10% of global annual turnover, whichever is the higher. It will be able to consider taking enforcement action, which may include business disruption measures, against any in-scope company worldwide that provides services to UK users. The government will reserve the right to introduce criminal sanctions for senior managers if they fail to comply with the regulator's information requests. Ofcom will take a proportionate approach to its enforcement activity. The government will establish a statutory appeals route that is accessible to companies.

39. The government will continue to assess the institutional landscape as its digital regulation programme progresses and will take action if necessary to ensure the landscape is coherent and streamlined.

What part will technology, education and awareness play in the solution?

Technology

40. The White Paper recognised the critical role of technology in improving user safety online, such as using artificial intelligence to identify harmful content quickly and accurately. The recent 'Safer Technology, Safer Users: The UK as a World-Leader in Safety Tech' report showed the UK is at the forefront of the rapidly developing safety tech industry, with the industry seeing an annual 35% growth rate since 2016.[23] The government will continue to invest in this sector, both to support companies in complying with the regime and to promote wider economic growth in the UK.

Safety by design, media literacy and engaging with information

41. Encouraging companies to build safer products and services will be key to delivering a successful regulatory regime. Our proposed safety by design framework will set out clear principles and practical guidance on how companies can design safer online products and services. The government, Ofcom and industry will also do more to equip users with the skills they need to keep themselves and others safe online, starting with the publication of an online media literacy strategy. This will build on Ofcom's existing media literacy work. The government and Ofcom will consider the links between service design and media literacy as part of this.

Next steps

42. The Online Safety Bill, which will give effect to the regulatory framework outlined in this document, will be ready in 2021. The government also expects the Law Commission to produce recommendations concerning the reform of the criminal offences relating to harmful online communications in early 2021. The Law Commission is currently consulting on its proposals for updating the criminal law in this area.[24] The government will consider, where appropriate, implementing the Law Commission's final recommendations through the Online Safety Bill.

43. As the new regulatory framework will be the first comprehensive approach to tackling online harms in the world, the Secretary of State for Digital, Culture, Media and Sport will undertake a review of the effectiveness of the regime 2-5 years after entry into force. The government will produce a report setting out findings from the review and conclusions about whether changes are necessary, which will then be laid in Parliament. Parliament will have an opportunity to debate the findings of the report.

[23] 'Safer technology, safer users: The UK as a world-leader in Safety Tech' UK Government, May 2020 (last viewed in November 2020)

[24] 'Harmful Online Communications: The Criminal Offences'' Law Commission, September 2020 (last viewed in November 2020)

Part 1: Who will the new regulatory framework apply to?

Summary

Consultation questions covered in Part 1:

- *Are proposals for the online platforms and services in scope of the regulatory framework a suitable basis for an effective and proportionate approach?*

- The new regulatory framework will apply to companies whose services host user-generated content or facilitate interaction between users, one or more of whom is based in the UK, as well as search engines. Services playing a functional role in enabling online activity will remain out of scope, as will business-to-business services.
- Exemptions will be applied where the risk of harm is sufficiently low that any regulatory requirements would be disproportionate. The government will exempt services used internally by organisations, services managed by educational institutions that are already subject to regulatory or inspection frameworks (or similar processes) that address online harm, email and telephony providers, and services with limited user functionality. Ofcom will take a risk-based and proportionate approach to its regulatory activity, focusing on companies whose services pose the biggest risk of harm.
- The government will put in place safeguards to ensure that media freedom is upheld. Content and articles produced and published by news services on their own sites do not constitute user-generated content and therefore fall outside the scope of legislation. Below-the-line comments on articles on news publishers' sites will be explicitly exempted from scope. In order to protect media freedom, legislation will include robust protections for journalistic content shared on in-scope services.
- The regulatory framework will apply to public communication channels, and services where users expect a greater degree of privacy, such as online instant messaging services and closed groups. The regulator will set out how companies can fulfil their duty of care in codes of practice, including what measures are likely to be appropriate in the context of private communications.

Services in scope

White Paper: *The White Paper set out that the regulatory framework will apply to companies that provide services or tools that allow, enable or facilitate users to share or discover user-generated content, or interact with each other online. It noted that regulatory requirements will need to be flexible, risk-based and proportionate. Search engines will be included in the scope of the regulatory framework.*

Consultation responses and stakeholder engagement: *There was broad support for the proposed approach. Many parties expressed a need for clarity around organisations in scope. There were calls to exclude business-to-business services due to the lower risk of harm on those services.*

Initial government response: *The initial government response confirmed that only a small proportion of UK businesses (estimated to account to less than 5%) are likely to fall within the scope of the regulatory framework. It also confirmed that business-to-business services will be out of scope of regulation.*

Final policy position: *The government will be maintaining a broad regulatory scope encompassing services that host user generated content and facilitate interaction between users, as well as search engines. The government also recognises that some businesses and services present a lower risk than others and that any approach must be proportionate to the level of risk and companies' capacity to address harm. Specific exemptions have been introduced for low-risk services. For example, reviews and comments by users on a company's website which relate directly to the company, its products and services, or any of the content it publishes, will be out of scope.*

1.1 As set out in the White Paper, the companies in scope of the regulatory framework will be defined by the types of services they provide. Companies[25] will fall into scope if their services:

(a) host user-generated content which can be accessed by users in the UK; and/or
(b) facilitate public or private online interaction between service users, one or more of whom is in the UK.

This covers a broad range of services, including (among others) social media services, consumer cloud storage sites, video sharing platforms, online forums, dating services, online instant messaging services, peer-to-peer services, video games which enable interaction with other users online, and online marketplaces.

1.2 Only companies with direct control over the content and activity on a service will be subject to the duty of care. This means that business-to-business services will remain outside the scope of the regulatory framework. It also means that services which play a functional role in enabling online activity will remain out of scope, including internet service providers, virtual private networks, browsers, web-hosting companies, content delivery service providers, device manufacturers, app stores, enterprise private networks and security software. However, such services will, where appropriate, be legally required to comply with the regulator as part of any business disruption enforcement measures (see Part 4 for further details).

[25] In this document the term 'company' is used to refer (where appropriate) to all entities providing in-scope services, including incorporated and unincorporated associations, partnerships and individuals.

Box 1: User-generated content and user interactions

Legal definitions of these concepts will be set out in the legislation; however, these will cover:

User-generated content

- digital content (including text, images and audio) produced, promoted, generated or shared by users of an online service
- content may be paid-for or free, time-limited or permanent. It must have the potential to be accessed, viewed, consumed or shared by people other than the original producer, promoter, generator or creator

User interaction

- any public or private online interaction between service users with potential to create and promote user-generated content
- interaction may be one-to-one or one-to-many and may involve means other than text, images and audio

In both cases, **'user'** refers to any individual, business or organisation (private or public) that puts content on a third-party online service. Users may be members, subscribers or visitors to the service, and may generate content or interact directly or through an intermediary, such as an automated tool or a bot.

1.3 Search engines will be included in scope of the regulatory framework. Search engines do not host user-generated content directly or facilitate interaction between users. However, there is evidence of harm occurring on these services, including facilitating easy access to child sexual exploitation and abuse content online. There are clear actions they can take to mitigate the risk of harm and they will be expected to put in place proportionate systems and processes to keep their users safe. This could include: removing known child sexual abuse images from their image search results; identifying keywords used to access illegal content; ensuring algorithms and predictive searches do not promote relevant illegal content; and protecting users online by signposting to resources and support. Given the distinct nature of search engines, legislation and codes of practice will include specific material for them. All regulatory requirements will be proportionate, and respect the key role of search engines in enabling access to information online.

1.4 The White Paper consulted on defining private communications, and what regulatory requirements should apply to them. It also said that companies would not be required to monitor for illegal content on these services in order to protect user privacy.

1.5 The regulatory framework will apply to both public communication channels and services where users expect a greater degree of privacy - for example online instant messaging services and closed social media groups. All companies in scope will be required to fulfil the duty of care by ensuring that they take reasonably practicable steps to tackle relevant illegal content, and protect children where they are likely to access their services. The regulator will set out how companies can fulfil their duty of care in codes of practice, including what measures are likely to be appropriate in the context of private communications. This could include steps to make services safer by design, such as limiting the ability for anonymous adults to contact children. The scale, severity and complexity of child sexual exploitation and abuse is particularly concerning, with private channels being exploited by offenders. In light of this, Part 2 sets out the circumstances in which the regulator will have the power to require companies to use automated technology to identify child sexual exploitation and abuse.

Voluntary best practice guidance for infrastructure service providers

Box 2: The government will produce voluntary best practice guidance for infrastructure service providers which is separate from the online harms regime.

- Infrastructure service providers still have a role to play in combatting the most serious harms such as child sexual exploitation and abuse.
- For example, Internet Watch Foundation and its Internet Service Provider partners blocked 8.8 million attempts to access child sexual abuse content from the UK in a month-long period earlier this year,[26] and web hosting providers are making tools to detect child sexual abuse content available to their customers.[27][28]
- In light of this, the government will produce voluntary best practice guidance for infrastructure service providers, setting out where their actions can help identify and prevent child sexual exploitation and abuse. This guidance will be separate from the online harms regime.

Exemptions

1.6 Many companies and representative groups expressed concerns through the consultation about low-risk businesses being captured in scope of the new framework. The COVID-19 pandemic has also placed unprecedented challenges on UK businesses. In response, a number of services will be exempt from the regulatory requirements. These exemptions apply to specific services, rather than entire companies. These exemptions are:

(a) *Business services.* Online services which are used internally by organisations - such as intranets, customer relationship management systems, enterprise cloud storage, productivity tools and enterprise conferencing software - will be excluded from scope. The risk of harm on these services is low, as the user base is limited and users tend to be verified and acting in a professional capacity. Organisations will already have policies in place for protecting users and managing disputes. Requiring them to comply with the legislation would be a disproportionate regulatory burden.

(b) *Online services managed by educational institutions, where those institutions are already subject to sufficient safeguarding duties or expectations.* This includes platforms used by teachers, students, parents and alumni to communicate and

[26] 'Millions of attempts to access child sexual abuse online during lockdown' Internet Watch Foundation, 2020 (last viewed in November 2020)

[27] 'Fighting the harmful content problem' Microsoft (viewed in November 2020)

[28] 'Announcing the CSAM Scanning Tool, Free for All Cloudflare Customers' Cloudflare, December 2019 (last viewed in November 2020)

collaborate. This is to avoid unnecessarily adding to any online safeguarding regulatory or inspection frameworks (or similar processes) already in place.

(c) *Email and telephony.* Email communication, voice-only calls and SMS/MMS remain outside the scope of legislation. It is not clear what intermediary steps providers could be expected to take to tackle harm on these services before needing to resort to monitoring communications, so imposing a duty of care would be disproportionate.

Low-risk functionality exemption

1.7 The legislation will exempt many low-risk businesses with limited functionality. It will exempt user comments on digital content provided that they are in relation to content directly published by a service. This will include reviews and comments on products and services directly delivered by a company, as well as 'below the line comments' on articles and blogs. This approach avoids imposing costs on businesses to familiarise themselves with the legislation when they are unlikely to have to take action to comply with the duty of care, given the low risk that this functionality poses to most users. It will also help to ensure the protection of media freedom and freedom of speech.

1.8 The online harms regulatory framework has been designed to reduce the burden on UK business by focussing on the areas that present the greatest risk of harm. The government estimates that, overall, fewer than 3% of UK businesses in total will be in regulatory scope following the new exemptions outlined above.[29] Ofcom, as the regulator, will also take a deliberately risk-based and proportionate approach to companies in scope, some of whose services will be low-risk.

1.9 Any exemption creates the potential for harm to be displaced from other services, particularly as technology and user behaviour evolve. The government will exempt these functionalities in a way which allows the Secretary of State for Digital, Culture, Media and Sport to bring them into scope, should evidence of the level of risk they pose change.

Journalism

White Paper: *The White Paper committed to ensuring protections for freedom of expression within the regulatory framework. Subsequently, Ministers confirmed that there would be strong protections for journalistic content. The Conservative and Unionist Party Manifesto 2019 reaffirmed the commitment to the protection of media freedom in the legislation[30].*

Consultation responses and stakeholder engagement: *There were calls to exclude journalistic content from scope, to protect freedom of expression and avoid negatively affecting the public's ability to access information or undermining quality news' media.*

[29] DCMS Online Harms research (externally commissioned), 2020, publication date tbc.

[30] 'The Conservative and Unionist Party Manifesto' The Conservative and Unionist Party, 2020 (last viewed in November 2020)

Final policy position: *Content and articles produced and published by news websites on their own sites, and below-the-line comments published on these sites, will not be in scope of legislation. In order to protect media freedom, legislation will include robust protections for journalistic content shared on in-scope services. The government is committed to defending the invaluable role of a free media and is clear that online safety measures must do this. The government will continue to engage with a range of stakeholders to develop these proposals.*

1.10 Freedom of expression is at the heart of the regulatory framework and there will be strong safeguards to ensure that media freedom is upheld. Content and articles produced and published by news services on their own sites do not constitute user-generated content and so are out of scope. The government recognises the importance of below-the-line comments for enabling reader engagement with the news. User comments below articles on news publishers' sites will be explicitly exempted from scope. This will be achieved via the low-risk functionality exemption (see above).

1.11 Journalistic content is shared across the internet, on social media, forums and other websites. Journalists use social media services to report directly to their audiences. This content is subject to in-scope services' existing content moderation processes. This can result in journalistic content being removed for vague reasons, with limited opportunities for appeal. Media stakeholders have raised concerns that regulation may result in increased takedowns of journalistic content.

1.12 In order to protect media freedom, legislation will include robust protections for journalistic content shared on in-scope services. The government will continue to engage with a wide range of stakeholders to develop proposals that protect the invaluable role of a free media and ensure that the UK is the safest place in the world to be online.

Advertising

Box 3: Online harms regulation and advertising

1. The online advertising ecosystem is complicated and includes services within and also beyond the scope of the online harms regulatory framework. Last year the Secretary of State for Digital, Culture, Media and Sport announced a review of the way that the online advertising market is regulated in the UK, which is being considered through the Online Advertising Programme. This programme of work, amongst other areas of focus, is identifying where regulatory gaps may exist and ensuring that advertising regulation answers the needs of the changing advertising marketplace. It will consider a full range of approaches, including support to help regulators meet the challenges posed by new advertising technologies and the potential for changes to the regulatory landscape.

2. As part of the Online Advertising Programme, the Department for Digital, Culture, Media and Sport will launch a public consultation on measures to enhance how online advertising is regulated in the UK in the first half of 2021. The consultation will

build on the call for evidence launched on this subject earlier this year and will consider options to enhance the regulation of advertising content and placement online.

3. Separately, as part of the government's new strategy 'Tackling obesity: empowering adults and children to live healthier lives', the government has committed to introducing a watershed ban on the advertising of foods that are high in fat, sugar and salt (HFSS) on broadcast TV, as well as further restrictions online. The strategy also announced that the government wanted to explore going further online. A consultation has been published on how a total HFSS advertising restriction online would be introduced, and the response to this and the previous 2019 consultation will be published in early 2021, setting out plans in more detail.

4. As the government considers further action on these issues, it will seek to avoid duplication between these areas, ahead of future regulatory requirements.

5. Nevertheless, some types of advertising will still fall in scope of the online harms regulatory framework. The definition of user-generated content will encompass organic and influencer adverts that appear on services in scope of the legislation. This includes images or text posted from users' accounts to promote a product, service or brand, and may or may not be paid for. As these are indistinguishable from other forms of user-generated content, it is therefore important, for clarity and consistency, that online harms safety systems and processes apply to these advertising posts.

6. The Advertising Standards Authority will remain responsible for overseeing the regulation of advertising. It will continue to regulate the content of individual adverts and advertisers' compliance with the advertising codes. Policy or political arguments - both online and offline - which can be rebutted by rival campaigners as part of the normal course of political debate are not regulated and the government does not support such regulation. It is a matter for voters to decide whether they consider materials to be accurate or not. The laws on defamation and the long-standing electoral offence of false statements about a candidate would also remain in place.

Part 2: What harmful content or activity will the new regulatory framework apply to, and what action will companies need to take?

Summary

Consultation questions covered in Part 2:

- ❖ *What further steps could be taken to ensure the regulator will act in a targeted and proportionate manner?*
- ❖ *In developing a definition for private communications, what criteria should be considered?*
- ❖ *What channels or forums that can be considered private should be in scope of the regulatory framework? What specific requirements might be appropriate to apply to private channels and forums in order to tackle online harms?*

- The legislation will set out a general definition of the harmful content and activity covered by the duty of care. This will include only content or activity which gives rise to a reasonably foreseeable risk of harm to individuals, and which has a significant impact on users or others. A limited number of priority categories of harmful content, posing the greatest risk to individuals, will be set out in secondary legislation.

- All companies in scope will be required to understand the risk of harm to individuals on their services, and to put in place appropriate systems and processes to improve user safety and monitor their effectiveness. The legislation will not change companies' liability for individual items of illegal content that meet the definition of harm. Instead it will require companies to ensure that their policies and processes are adequate to protect their users.

- Recognising the importance of freedom of expression, the government will establish differentiated obligations on companies in scope with regard to different categories of content and activity. Only a small number of high-risk, high-reach Category 1 services will have to address legal but harmful content and activity accessed by adults on their services.

- The regulator will issue codes of practice to outline the systems and processes that companies can adopt to fulfil the duty of care, including what measures are likely to be appropriate in the context of private communications. The government is publishing interim codes on terrorism and child exploitation and sexual abuse alongside this document.

- The duty of care will apply to disinformation and misinformation that could cause harm to individuals, such as anti-vaccination content. The legislation will introduce additional provisions targeted at building understanding and driving action to tackle disinformation and misinformation. These provisions will include an expert working group which will build consensus and technical knowledge on how to tackle disinformation and misinformation.

Definition of harm

White Paper: *The White Paper set out an initial list of harms in scope but made clear this was, by design, neither exhaustive nor fixed. A static list could prevent swift regulatory action to address new forms and types of online harm. It also set out specific exclusions from scope where there are existing government initiatives to tackle these harms.*

Consultation responses and stakeholder engagement: *Stakeholders wanted more detail on the breadth of both services and harms in scope. There were calls to protect freedom of expression and a focus on protecting children. Some suggested that further work should be done to increase education and public awareness of online harms.*

Final policy position: *The legislation will set out a general definition of the harmful content and activity in scope of the regime. A limited number of priority categories of harmful content will be set out in secondary legislation. Some categories of harmful content will be explicitly excluded, to avoid regulatory duplication. This will provide legal certainty for companies and users and prioritise action on the biggest threats of harm.*

Harmful content and activity covered by the duty of care

2.1 The regulatory framework will require companies to have effective systems and processes in place to improve user safety. The response to the consultation flagged concerns about the broad range of potential harms in scope of the regime and called for greater clarity. The legislation will set out a general definition of the harmful content and activity in scope. This will help provide legal certainty for companies and users and set a clearly defined statutory remit for Ofcom.

2.2 The legislation will set out that online content and activity should be considered harmful, and therefore in scope of the regime, where it gives rise to a reasonably foreseeable risk of a significant adverse physical or psychological impact on individuals. Companies will not have to address content or activity which does not pose a reasonably foreseeable risk of harm, or which has a minor impact on users or others. Harms to organisations will not be in scope of the regime.

2.3 A limited number of priority categories of harmful content, posing the greatest risk to users, will be set out in secondary legislation. These will cover (i) priority categories of criminal offences (including child sexual exploitation and abuse, terrorism, hate crime and sale of illegal drugs and weapons) (ii) priority categories of harmful content and activity affecting children, such as pornography or violent content, and (iii) priority categories of harmful content and activity that is legal when accessed by adults, but which may be harmful to them, such as abuse and content about eating disorders, self-harm or suicide. Further information on the approach, and the expectations on companies, is set out below.

2.4 In line with the position set out in the White Paper, a number of harms will be excluded from scope where there are existing legislative, regulatory and other governmental initiatives in place. The following will be excluded from scope:

- Harms resulting from breaches of intellectual property rights;
- Harms resulting from breaches of data protection legislation;
- Harms resulting from fraud;
- Harms resulting from breaches of consumer protection law;
- Harms resulting from cyber security breaches or hacking.

The online harms regulatory framework will not aim to tackle harm occurring through the dark web.[31] A law enforcement response to tackle criminal activity on the dark web is more suitable than a regulatory approach.

Box 4: Online Fraud and Sale of Unsafe Goods

White Paper: *The White Paper did not set out a definitive position on whether economic and financial harms to individuals, including online fraud and sale of unsafe goods, would be in scope of the new regulatory framework.*

Consultation responses and stakeholder engagement: *A number of organisations suggested that economic harms (for instance, online fraud) should be in scope, noting that such activity could also lead to significant psychological harm. Others argued that the scope of the regulatory framework was too broad, and that any further extension would pose disproportionate regulatory burdens on businesses.*

Final policy position: *The government is deeply concerned by the growth, impact and scale of online fraud, recognising the devastating harm these types of fraud can cause.The government has determined that the fraud threat will be most effectively tackled by other mechanisms and as such the legislation will not require companies to tackle online fraud. We are working closely with industry, regulators and consumer groups to consider additional legislative and non-legislative solutions. This ongoing programme of work aims to effectively address the harms posed by all elements of online fraud in a cohesive and robust way. This includes work on the Online Advertising Programme, led by the Department for Digital, Culture, Media and Sport, which will be considering further regulation of online advertising to reduce online harms, including fraud.*

As noted elsewhere, most forms of advertising, fake websites and data and cyber-security breaches are not in scope of the online harms regulatory framework. This would have limited the impact the regulatory framework would have had on tackling fraud if it were in scope.
The government is committed to tackling the sale of unsafe consumer products. The Office for Product Safety and Standards has a clear remit for consumer product safety, including products sold online. In order to avoid regulatory duplication the sale of unsafe products will be excluded from the online harms regulatory framework.

[31] The dark web is made up of a number of untraceable online websites. Specific software and search engines must be used to access the websites.

Duty of care and principles of the regulatory framework

White Paper: *The White Paper stated that there would be a new statutory duty of care to make companies take more responsibility for the safety of their users. This duty would be risk-based and proportionate and focused on systems and processes, not individual pieces of content. Important principles would apply to the regulatory framework including users' rights to freedom of expression and privacy, innovation and protecting small and medium-sized enterprises.*

Consultation responses and stakeholder engagement: *Many stakeholders welcomed the approach, noting that this would underpin an effective, future-proofed framework. Nevertheless, industry responses sought greater reassurance and certainty about how it would be proportionate in practice, particularly for small and medium-sized enterprises; and how flexibility would be balanced with certainty about what the duty of care requires of companies. Rights groups and industry also emphasised the need to provide more certainty about how safety would be balanced with freedom of expression, particularly in relation to legal but harmful content.*

Final policy position: *In order to provide more clarity and target effectiveness, the duty of care has been refined. It will cover content and activity that could cause harm to individuals. The legislation will also introduce additional provisions targeted at building understanding and driving action to tackle disinformation and misinformation.*

2.5 The primary purpose of the duty of care will be to improve safety for users of online services, and to prevent other people from being harmed as a direct consequence of content or activity on those services.

How the duty of care works

2.6 The duty of care consists of two parts. The first part relates to the duties on companies and the second part relates to the regulator's duties and functions. Companies and the regulator will be required to carry out their responsibilities under the framework in line with a range of guiding principles (not all will apply to both). Further details on how the regulatory framework will be delivered against the guiding principles are set out in **Annex A**.

Duties on companies in scope

2.7 The primary responsibility for each company in scope will be to take action to prevent user-generated content or activity on their services causing significant physical or psychological harm to individuals. To do this they will complete an assessment of the risks associated with their services and take reasonable steps to reduce the risks of harms they have identified occuring.

2.8 The steps a company needs to take will depend, for example, on the risk and severity of harm occurring, the number, age and profile of their users and the company's size. Search engines will need to assess the risk of harm occurring across their entire service. Ofcom will provide guidance specific to search engines regarding regulatory expectations.

2.9 Companies will fulfil their duty of care by putting in place systems and processes that improve user safety on their services. These systems and processes will include, for example, user tools, content moderation and recommendation procedures. The proposed safety by design framework (detailed in Part 5) will support companies to understand how they can improve user safety through safer service and product design choices.

2.10 Robust protections for freedom of expression have been built into the design of duties on companies. Companies will be required to consider users' rights, including freedom of expression online, both as part of their risk assessments and when they make decisions on what safety systems and processes to put in place on their services. Regulation will ensure transparent and consistent application of companies' terms and conditions relating to harmful content. This will both empower adult users to keep themselves safe online, and protect freedom of expression by preventing companies from arbitrarily removing content.

2.11 The regulatory framework will improve user safety online but it will not eliminate harm or the risk of harm entirely. Users must be able to report harm when it does occur and seek redress. They must also be able to challenge wrongful takedown and raise concerns about companies' compliance with their duties. This is essential to improving users' safety, and to help companies understand the risk and incidence of harm on their services.

2.12 All companies in scope will have a specific legal duty to have effective and accessible reporting and redress mechanisms. This will cover harmful content and activity, infringement of rights (such as over-takedown), or broader concerns about a company's compliance with its regulatory duties. Ofcom's codes of practice will set out expectations for these mechanisms. The government expects the codes to cover areas such as accessibility (including to children), transparency, communication with users, signposting and appeals. Expectations on companies will be risk-based and proportionate, and will correspond to the types of content and activity which different services are required to address. For example, the smallest and lowest risk companies might need to give only a contact email address, while larger companies offering higher-risk functionalities will be expected to provide a fuller suite of measures.

2.13 The government will not mandate specific forms of redress, and companies will not be required to provide financial compensation to users (other than in accordance with any existing legal liability). Forms of redress offered by companies could include: content removal; sanctions against offending users; reversal of wrongful content removal or sanctions; mediation; or changes to company processes and policies.

2.14 The regulatory framework will not establish new avenues for individuals to sue companies. However, the existing legal rights individuals have to bring actions against companies will not be affected. As outlined in the White Paper, the government expects legal action to become more accessible to users as the evidence base around online harms grows, and as regulatory precedent is established. Users will be able to use regulatory decisions that are publicly available as evidence in any relevant legal action they pursue.

Box 5: Service design and the risk of online harms

- The design of a service and its features can be one of the factors that contributes to the risk of harm occurring to a user. For example, a service is likely to be higher risk if it has features such as: allowing children to be contacted by unknown adult users; allowing all users - including children - to live-stream themselves; and including private messaging channels where the content on those private channels is not or cannot be moderated. A lower risk service might include features such as: the ability to moderate all content; having public messaging forums with text content only; and taking steps to ensure an age appropriate environment for children, for example by restricting contact of children by unknown users.

- As part of their duty of care, companies in scope will be expected to consider, as part of their regular risk assessments, the risk of online harms posed by their service, including the risk presented by the design of their service and its features. Companies will be expected to reassess the risk of online harms if they are planning significant changes to their services.

- Following the risk assessment, companies will be required to take steps to address the risks they have identified. This will be key to them fulfilling their duty of care to their users and delivering a higher level of protection for children.

- The regulator will set out the steps that companies should take to address the risk posed by their services, and ultimately will have the power to assess whether the steps taken are sufficient to fulfil the company's regulatory requirements. Failure to fulfil the duty of care may result in the regulator taking robust enforcement action.

- The decisions taken by a company on the design or functionality of their service will not exempt them from needing to comply with other regulatory requirements. For example, all companies in scope must comply with information requests from the regulator. In tightly prescribed circumstances, and subject to stringent legal safeguards, the regulator will be able to require the use of highly accurate technology to identify specific categories of illegal child sexual abuse or terrorist content and activity. As with all regulatory requirements, the onus will be on the company to comply with these requirements.

Differentiated expectations on companies

White Paper: *The White Paper set out that all services in scope will be required to address illegal and legal but harmful content and activity. It stated that the regulatory approach would impose more specific and stringent requirements for illegal harms than for content and activity which are legal but have the potential to cause harm, depending on the context. It acknowledged that the impact of harmful content and activity can be particularly damaging for children and placed particular emphasis on keeping children safe online.*

Consultation responses and stakeholder engagement: *The consultation responses flagged concerns about the broad scope of harms, calling for greater clarity and highlighting the subjectivity inherent in identifying many of the harms, especially those which are legal. Many respondents objected to the latter being in scope. There were concerns that proposals could impact freedom of expression online. Respondents to the consultation welcomed the approach to the protection of children.*

Final policy position: *The initial government response developed the original position, confirming a differentiated approach for illegal content and activity versus content that is legal but harmful. Only companies providing Category 1 services will have to take action in respect of adult users accessing legal but harmful content on their services.*

All companies in scope will be expected to assess whether children are likely to access their services, and if so, take measures to protect children on their services including reasonable steps to prevent them from accessing age-inappropriate and harmful content.

2.15 The regulatory framework will establish differentiated expectations on companies in scope with regard to different types of content and activity. This will ensure companies prioritise tackling relevant illegal content and activity on their services, and that children are protected from age-inappropriate and harmful content online. The differentiated approach can be summarised as follows:

- All companies will be required to take action with regard to relevant illegal content and activity.
- All companies will be required to assess the likelihood of children accessing their services. If they assess that children are likely to access their services, they will be required to provide additional protections for children using them.
- Only companies with Category 1 services will be required to take action with regard to legal but harmful content and activity accessed by adults. This is because services offering extensive functions for sharing content and interacting with large numbers of users pose a significantly increased risk of harm from legal but harmful content. The approach will protect freedom of expression and mitigate the risk of disproportionate burdens on small businesses. It will also address the current mismatch between companies' stated safety policies and many users' experiences online which, due to their scale, is a particular challenge on the largest social media services.

Designating Category 1 services

2.16 Category 1 services will be determined through a three-step process. First, the primary legislation will set out high level factors which lead to significant risk of harm occurring to adults through legal but harmful content. These factors will be: the size of a service's audience (because harm is more likely to occur on services with larger user bases, for example due to rapid spread of content and 'pile-on' abuse); and the functionalities it offers (because certain functionalities, such as the ability to share content widely or contact users anonymously, are more likely to give rise to harm).

2.17 Second, the government will determine and publish thresholds for each of the factors. Ofcom will be required to provide non-binding advice to the government on where these thresholds should be set. The final decision on thresholds will lie with the government, to ensure democratic oversight of the scope of the regulatory framework.

2.18 Ofcom will then be required to assess services against these thresholds and publish a register of all those which meet both thresholds. These services will be designated as Category 1 services and be required to take action against legal but harmful content accessed by adults. Ofcom will be able to add services to the list of Category 1 services if they reach the thresholds, and to remove services if they no longer meet the thresholds. If a company believes its service has wrongly been designated as Category 1, then it will be able to appeal to an appropriate tribunal (further detail on Appeals is set out in Part 4). Ofcom will also be able to provide advice to the government if it considers a change to the thresholds to be necessary.

Illegal content and activity

2.19 All companies in scope will need to take action to prevent the use of their services for criminal activity. They will need to ensure that illegal content is removed expeditiously and that the risk of it appearing and spreading across their services is minimised by effective systems.

2.20 The government will set priority categories of offences in secondary legislation, against which companies will be required to take particularly robust action. These will be offences posing the greatest risk of harm, taking account of the number of people likely to be affected and how severely they might be harmed. Examples of priority categories of offences include child sexual exploitation and abuse and terrorism. The identification of priority categories of offences will focus companies', and the regulator's, efforts on the most harmful issues. Companies will still be required to tackle other relevant illegal material on their services, where this is identified through their systems or where it is reported to them.

2.21 For priority categories of offences, companies will need to consider, based on a risk assessment, what systems and processes are necessary to identify, assess and address such offences (for example devoting more resources to content moderation or limiting algorithmic promotion of content). Recognising the severity of child sexual exploitation and abuse and terrorism, companies may be required to proactively identify and block or remove this type of illegal material if other steps have not been effective and safeguards are in place. Further details are set out later on in Part 2.

2.22 All companies in scope must additionally take steps to minimise the risk of other relevant illegal content and activity occurring on their services. This will require putting in place effective user reporting and redress mechanisms for dealing with such illegal content and activity.

2.23 Companies may already be liable for illegal content and activity on their services. Under existing law, they may be liable for such content if they have been notified of its existence, have subsequently failed to remove it in good time, and the hosting of such content gives rise to criminal or civil liability. These existing legal responsibilities will remain in place.

2.24 The regulatory framework will require companies to address illegal content and activity which could constitute a UK criminal offence or an element of a UK criminal offence and which meets the definition of harm, as set out above. It will not cover online material which only gives rise to a risk of civil liability (e.g. negligence or defamation). Some areas of criminal law will be excluded, as set out above in paragraph 2.4**.**

Freedom of expression and relevant illegal material

2.25 To avoid companies taking an overly risk-averse approach to the identification and removal of material likely to be illegal, the regulatory framework will enshrine strong safeguards for freedom of expression. Further details are included in Annex A. Companies will be required to consider the impact on and safeguards for users' rights when designing and deploying content moderation systems and processes. This might involve engaging with stakeholders in the development of their content moderation policies, considering the use of appropriate automated tools, and ensuring appropriate training for human moderators. Companies should also take reasonable steps to monitor and evaluate the effectiveness of their systems, including considering the amount of legitimate content that was incorrectly removed.

2.26 The regulatory framework will also require companies to give users a right to challenge content removal, as an important protection for freedom of expression. Certain companies will also need to produce transparency reports, which are likely to include information about their measures to uphold freedom of expression and privacy (see Part 4 for more information on transparency).

2.27 The online harms regime will not change companies' liability for individual items of illegal content that meet the definition of harm. Instead it will require companies to ensure that their policies and processes are adequate to protect their users. Where moderation procedures meet the above objectives, individual instances of illegal content or activity appearing on a company's services will not necessarily mean it has failed to fulfil the duty of care.

Box 6: Taking, making and sharing intimate images without consent

- The evolution of technology has made it easier for users to create images to send to friends, family or post en masse to the public. It also means that it is easier to distribute images of individuals without consent. This is particularly harmful when those images are 'intimate' in nature, such as revenge and deepfake pornography.

- Currently, there is no single criminal offence in England and Wales that captures the taking, making and sharing of intimate images without consent. Instead, we have a range of offences that have developed over time, some of which existed before the rise of the internet and use of smartphones.

- To ensure that legislation provides victims with the right support and protection from these harmful behaviours, the Ministry of Justice has sponsored the Law Commission to review the law around the taking, making and sharing of non-consensual intimate images. The Law Commission has not yet issued its draft recommendations for this review but following the final recommendations the government will consider taking forward the proposals, where appropriate, in a legislative vehicle.

- All companies in scope of the duty of care will be required to take action against illegal content and activity, including intimate image abuse.

Legal but harmful content and activity accessed by adults

2.28 Only companies providing Category 1 services will be expected to take steps in respect of legal but harmful content and activity that is accessed by adults. The legislation will not require the removal of specific pieces of legal content. Companies must consider the impacts of their decisions regarding moderation and design choices on user safety. The approach will ensure transparent and consistent application of companies' terms and conditions relating to harmful content. This will both empower adult users to keep themselves safe online and protect freedom of expression, by preventing companies from arbitrarily removing content.

2.29 The government will set out priority categories of legal but harmful material in secondary legislation (e.g. content promoting self-harm, hate content, online abuse that does not meet the threshold of a criminal offence, and content encouraging or promoting eating disorders). Ofcom will be required to provide non-binding advice to the government on what should be included in that secondary legislation. Categories of legal but harmful material must meet the definition of harmful content and activity described in paragraph 2.2. This approach will ensure that the regulatory framework provides sufficient clarity for businesses, users and the regulator about the categories of legal but harmful material that these companies should, at a minimum, address through their terms and conditions.

Box 7: Material that, of itself, may not be illegal but is linked to child sexual exploitation and abuse online

- The government remains committed to taking action against material that may not be illegal, but is linked to child sexual exploitation and abuse online. Such material can have a devastating impact on victims, contributing to their re-traumatisation and facilitating further offences.

- The government has engaged extensively with tech companies on the importance of responding to this content. In March 2020, the Voluntary Principles to Counter Online Child Sexual Exploitation and Abuse were launched, endorsed by a range of tech companies and the UK, US, Canadian, Australian and New Zealand governments.[32] These recognise the importance of taking appropriate action on certain images, videos, discussions and other material which may fall below the threshold of illegal but still warrant action. The government will continue to explore regulatory and legal options to ensure companies are taking effective and consistent action to tackle this content.

2.30 Companies providing Category 1 services will be required to undertake regular risk assessments to identify legal but harmful material on these services, covering both the priority

[32] '[Voluntary Principles to Counter Online Child Sexual Exploitation and Abuse](#)' Five Country Ministerial (last viewed in November 2020)

categories set out in secondary legislation and any other types of harm present or at risk of arising. Risk assessments should consider the risk to adult users, including vulnerable users. Companies providing Category 1 services will use the definition of harmful material in paragraph 2.2 to identify and notify the regulator of emerging legal but harmful harms. The regulator's codes of practice will include information on the risk assessment process.

2.31 These companies will be required to set clear and accessible terms and conditions which explicitly state how they will handle the priority categories of legal but harmful material established in legislation, and any others identified by them through their risk assessment. They will need to make clear to users what is acceptable on their services for such content, and how it will be treated across their services. Companies will be expected to consult with civil society and expert groups when developing their terms and conditions. This will encourage the adoption of terms and conditions that meet user needs and build on existing best practice on how to effectively tackle different types of harmful content and activity.

2.32 These terms and conditions must be enforced consistently and transparently, irrespective of what the company's policy is. This will include having effective and accessible reporting and redress mechanisms, with the regulator's codes of practice setting out the steps which companies can take to meet expectations. Terms and conditions will not simply be about accepting or removing content. They could include, for example, circumstances in which content has a label applied to it, or is de-prioritised. They could also include circumstances in which users are signposted towards support, or nudged in order to discourage behaviour.

2.33 This approach will empower adult users to keep themselves safe online, while ensuring that the legislation will not require companies providing Category 1 services to remove specific pieces of legal content unless specified as not permitted by their terms and conditions. It will be particularly beneficial for vulnerable adults and those disproportionately affected by online harms, including groups with protected characteristics or those with particular mental or physical health conditions, as they are currently more likely to experience harm associated with such content or activity online. It will also ensure companies providing Category 1 services are accountable for their public commitments in their terms and conditions.

2.34 This approach recognises the importance of high risk, high reach platforms as public forums where people can engage in robust debate online. Companies will not be able to arbitrarily remove controversial viewpoints and users will be able to seek redress if they feel content has been removed unfairly. When combined with transparency requirements (see Part 4), the duty to consistently apply terms and conditions will also increase understanding about what content is taken down and why. In this way, regulation will promote and safeguard pluralism online, while ensuring companies can be held to account for their commitments to uphold freedom of expression.

Box 8: Safety by design

- The White Paper recognised that companies themselves have a crucial role to play in tackling the proliferation of online harms. The design of an online product or service can give rise to harm or help protect against it.

- The government's forthcoming safety by design framework will set out what 'good' looks like for safe product and service design. The framework will be open source and developed with industry, subject and technical experts. It will contain clear principles and practical guidance for product designers, managers and developers on how to build safer online products and services from the outset. Further details are in Part 5.

- The safety by design framework will be an important step in ensuring that all companies, especially small businesses, are equipped with the know-how to effectively embed safety into the design of their online products and services, to help minimise regulatory burdens and support fulfilment of the duty of care.

- Device security also has an important role in user safety. In January 2020, the Minister for Digital Infrastructure announced that the government would be developing legislation to protect citizens and the wider economy from the harms that can arise from 'smart', Internet of Things (IoT) or 'internet-connected' devices that lack important cyber-security measures. This work is underway with a view to introducing legislation as soon as parliamentary time becomes available.

Box 9: Anonymous Abuse

- As set out in the White Paper, anonymous abuse can have a significant impact on victims, whether members of the public or high-profile public figures. It is important that the regulatory framework adequately addresses this issue, whilst protecting freedom of expression.

- Anonymous abuse has been on the rise. In a sample of 4.2 million tweets collected during the 2019 General Election campaign, abusive replies sent to candidates were found in nearly 4.5% of all replies, compared to just under 3.3% in the 2017 General Election.[33]

- The consultation did not specifically cover anonymous abuse but respondents put forward arguments both for and against preserving online anonymity, particularly in regard to protecting the identity of those individuals who flag harmful content.

- The regulatory framework will address abuse online, including anonymous abuse, whilst protecting freedom of expression and the legitimate use of anonymity online by groups such as human rights advocates, whistleblowers and survivors of abuse. The legislation will, therefore, not put any new limits on online anonymity.

- Under the duty of care, all companies in scope will be expected to address anonymous online abuse that is illegal through effective systems and processes. Where companies providing Category 1 services prohibit legal but harmful online

[33] 'Online Abuse toward Candidates during the UK General Election 2019: Working Paper' Gorell and others, January 2020 (last viewed in November 2020)

abuse, they will need to ensure their terms and conditions are clear about how this applies to abuse perpetrated anonymously. They will then need to enforce these terms and conditions consistently and transparently.

- Being anonymous online does not give anyone the right to abuse others. The police have a range of legal powers to identify individuals who attempt to use anonymity to escape sanctions for online abuse, where the activity is illegal. The government is continuing to review with law enforcement whether the current powers are sufficient to tackle illegal anonymous abuse online. The outcome of that work will inform the government's future position in relation to illegal anonymous abuse online.

- The government recognises that in the context of online abuse, the line between illegal and legal behaviour is not well understood. The Law Commission has reviewed the legal framework relating to abusive and offensive communications online. They are now consulting on their provisional proposals, which aim to improve the existing communications offences, ensuring the law is clearer and more effectively targets serious harm online.[34]

- As highlighted in their consultation, the Commission acknowledges that anonymity online often facilitates and encourages abusive behaviours. Combined with an online disinhibition effect, abusive behaviours, such as pile-on harassment, are much easier to engage in on a practical level.

- To deal with such abusive behaviours online, the Commission has put forward several recommendations. These include replacing existing offences with new laws which more effectively criminalise online behaviours likely to cause harm. These proposals are subject to consultation.

- The Law Commission is expected to provide its recommendations for reform of the criminal law in this area in early 2021. Once the final recommendations have been published the government will consider, where appropriate, whether to bring these recommendations into law as part of the Online Safety Bill.

- Intimidation and abuse in public life can also stop talented individuals, particularly women and those from minority backgrounds, from standing for public office, or undertaking high profile roles such as journalism. Journalists are often subject to online abuse and harassment, which can undermine their ability to carry out their vital democratic function.

- The government is therefore taking forward a co-ordinated programme of work to safeguard the integrity and security of our democratic processes. Under the Defending Democracy programme, a key priority is tackling the intimidation of elected officials by strengthening our legislative framework, driving policy across government, and engaging with partners.

[34] 'Harmful Online Communications: The Criminal Offences'' Law Commission, September 2020 (last viewed in November 2020)

- We are also encouraging respect for open, fair and safe democratic participation for voters and candidates by implementing the recommendations set out in the government's response to the Committee on Standards in Public Life report on Intimidation in Public Life.

- The government has committed to introduce a new sanction of intimidation against candidates or campaigners, either in person or online. The new electoral sanction is being developed to crack down on the intimidation and abuse being suffered by those at the forefront of public service. The government has also committed to the development of a National Action Plan for the Safety of Journalists.

Content and activity that is legal but harmful to children

2.35 The online harms regime will ensure the most comprehensive approach possible to protecting children. It will deliver the objectives of Part 3 of the Digital Economy Act, to protect children from accessing online pornography, and go further to protect children from a broader range of harmful and age-inappropriate content on all services in scope.

2.36 The framework will deliver a higher level of protection for children than for adults. All companies in scope will be required to assess the likelihood of children accessing their services. Only services which are likely to be accessed by children will be required to provide additional protections for children using them. This is the approach taken in the Information Commissioner's Age Appropriate Design Code, which requires companies to apply the Code's standards for protecting children's personal data where they have assessed that children are 'likely to access' their service. This will provide consistency for companies who may be required to comply with both the Age Appropriate Design Code and the duty of care.

2.37 Companies which have assessed their service as likely to be accessed by children will be required to conduct a child safety risk assessment of their service specifically for children, identify and implement proportionate mitigations to protect children, and monitor these for effectiveness. Companies will be required to undertake regular child safety risk assessments to identify legal but harmful material on their services impacting children, covering both the priority categories set out in secondary legislation (as detailed in paragraph 2.38 below) and any other types of harm present or at risk of arising to children. These companies will also be required to assess the risks that material on their services poses to children of different ages and to put in place age-appropriate protective measures. The regulator will be required to have regard to the fact that children have different needs at different ages when preparing codes of practice relevant to the protection of children. The regulator's codes of practice will include guidance on the risk assessment process. Companies will also need to put in place effective and accessible user reporting and redress mechanisms for content and activity which is harmful to children.

2.38 In addition to the approach for priority categories for illegal material and legal but harmful material accessed by adults described above, the government will also set out in secondary legislation priority categories of legal but harmful content and activity impacting children. The regulator will be required to provide non-binding advice to the government on what should be included in those categories. Categories of legal but harmful material impacting children must meet the definition of harmful content and activity described in paragraph 2.2. This approach will ensure that the regulatory framework provides sufficient clarity for businesses, users and the regulator about the categories of legal but harmful material impacting children that companies in scope should, at a minimum, take action on.

2.39 Companies with services likely to be accessed by children will need to make clear what is acceptable on their services for legal but harmful material as described above for adults. The regulator will determine appropriate levels of risk-based and proportionate protection for children and set out through its codes of practice the steps companies need to take. This is expected to cover legal but harmful content and activity such as cyberbullying, and access to age-inappropriate content such as online pornography. Specific measures required to address illegal harms such as child sexual exploitation and abuse are covered above in paragraphs 2.19 to 2.24.

2.40 The regulator will focus on ensuring that companies whose services are likely to be accessed by children have good systems and processes in place to protect children. This includes providing terms and conditions and user redress mechanisms that are suitable for children as well as more transparency about how services are providing greater protection.

2.41 Under our proposals companies will be expected to use a range of tools proportionately, to take reasonable steps to prevent children from accessing age-inappropriate content and to protect them from other harms. This includes, for example, the use of age assurance and age verification technologies, which are expected to play a key role for companies in order to fulfil their duty of care.

2.42 The government would not in every case expect age assurance technologies to be used to block children from content or services, but where appropriate, to protect children within a service and enhance a child user's experience by tailoring safety features to the age of the user. For example, the Lego Life app requires parental consent to unlock features and functions, to provide an age-appropriate service. The proposed safety by design framework will also reflect these design objectives in its guidance.

2.43 Although the government will not be mandating the use of specific technological approaches through the legislation to prevent children from accessing age-inappropriate content and to protect them from other harms, the government does expect that the regulatory framework will drive innovation and take-up of age assurance and, where appropriate, age verification technologies. The government is working closely with stakeholders across industry to establish the right conditions for the market to deliver these technical solutions ahead of the legislative requirements coming into force.

2.44 Technical standards also have an important role to play in tackling online harms. In line with this approach, the Department for Digital, Culture, Media and Sport is supporting the update of Publicly Available Standard 1296: 2018 ‘online age checking’. The Department for Digital, Culture, Media and Sport recognises the benefit PAS 1296 brings to the age assurance sector and to child online safety. The Department has contributed funding and is working closely with the British Standards Institute and other relevant stakeholders to bring the standard in line with current policy and industry needs.

2.45 In keeping with its existing priorities in broadcasting, the government expects Ofcom to prioritise children in its approach to enforcement in accordance with the principle of delivering a higher level of protection to children. In its enforcement guidelines, Ofcom will be required to set out how it will take into account any impact on children due to a company's failure to fulfil its duty of care.

Box 10: Online Harms and the Digital Economy Act

- In October 2019 the government announced that it would deliver the objective of protecting children online through the online harms regulatory framework instead of Part 3 of the Digital Economy Act 2017. The government has carefully reviewed how to ensure the objectives of the Digital Economy Act will be delivered by the framework. Through the regulatory framework, the government will go further to protect children from a broader range of harmful and age-inappropriate content, across a wider range of sites in scope, going beyond the Digital Economy Act's focus on online pornography on commercial adult sites.

- One of the criticisms of the Digital Economy Act was that its scope did not cover social media companies where a considerable quantity of pornographic material is accessible to children. The government's new approach will include social media companies and sites where user-generated content can be widely shared, including commercial pornography sites. Where pornography sites have such functionalities (including video and image sharing, commenting and live streaming) they will be subject to the duty of care. The online harms regime will capture both the most visited pornography sites and pornography on social media, therefore covering the vast majority of sites where children are exposed to pornography. Taken together we expect this to bring into scope more online pornography that children can currently access than the narrower scope of the Digital Economy Act.

- The regulator will determine appropriate levels of risk-based and proportionate protection for children. Companies in scope which are likely to be accessed by children will need to put in place measures to keep children safe from harmful activity and prevent them from accessing age-inappropriate or harmful content, including online pornography.

- The online harms legislation will not mandate the use of specific technological approaches to prevent children from accessing age-inappropriate content and to protect them from other harms. However, the government expects the regulator will take a robust approach to sites that pose the highest risk of harm to children, including sites hosting online pornography. This may include recommending the use of age assurance or verification technologies.

Interim measures from the government

2.46 The government is already undertaking initiatives to keep children and young people safe online and to build momentum ahead of the implementation of the online harms regime.

2.47 These measures include providing practical guidance for business on how to improve child safety online, steps companies can take to tackle cyberbullying, collaboration between government and industry to better understand the impacts of online harms on users, and cross-government research on child safety.

Box 11: Interim measures ahead of the online harms regulatory framework

'One Stop Shop for Companies on Protecting Children Online'

- The government will publish a 'One Stop Shop' with practical guidance for companies on how to protect children online. It will be designed as an interim tool to support businesses ahead of the regulatory framework.

- The One Stop Shop will support smaller companies in particular, providing practical advice to help them better understand child online harms and their existing legal requirements.

Cyberbullying

- The consultation highlighted that vulnerable young people are in particular need of support to stay safe online to tackle harms such as cyberbullying. We expect the regulator to set out steps companies need to take to tackle cyberbullying in its codes of practice. The Social Media Code of Practice, published alongside the White Paper, sets out the principles that companies should adhere to in the interim before the regulator is operational.

- In the longer term, the government will align its work on cyberbullying with the cross-government plan on tackling loneliness, recognising that loneliness, particularly amongst young people, can be exacerbated or directly caused by cyberbullying. The government will conduct further research and develop further guidance on tackling cyberbullying as part of this.

Screen time

- Being online can be a hugely positive experience for children and young people. However, the impact of harmful content and activity can be particularly damaging for children and there is also growing concern about the relationship between social media and the mental health of children and young people.

- In 2019, the UK Chief Medical Officers conducted a systematic evidence review on children and young people's screen and social media use. Whilst the research did not present evidence of a causal relationship between screen-based activities and mental health problems, it did find some associations between screen-based activities and negative effects, such as increased risk of anxiety or depression. The Chief Medical Officers therefore advised a precautionary approach to screen time, including agreeing boundaries with children and young people around their screen usage and considering the impact that screen use has on health promoting activities such as sleep.

- Since the Chief Medical Officers' review, the Department for Health and Social Care has commissioned research to explore the views of children and young people to help prioritise research questions on social media and mental health. It will also be developing robust methodologies to better examine the relationship between the two.

Box 12: Collaboration with industry

Harmful content including suicide, self-harm, and eating disorder content:

- The online harms framework will place regulatory responsibilities on in-scope companies likely to be accessed by children to protect their child users from harmful content and activity, including suicide, self-harm and eating disorder content. However there are wider government-led initiatives to develop voluntary cooperation in this area ahead of legislation.

- The Department for Health and Social Care has coordinated a strategic partnership with social media companies and the Samaritans to set guidance on moderating suicide and self-harm content, and educating users to stay safe online.

Box 13: Cross-government research

Verification of Children Online

- Companies will need to know which of their users are children and this is likely to be achieved through the use of age assurance technologies.

- Age assurance is the broad term given to the spectrum of measures that can be used to assure a user's age online. Age assurance allows companies and users to jointly choose from a range of measures that are appropriate to the specific risks posed and their service needs. The selected methods may rely on different sources of data, which may have different privacy implications and cost models.

- The Department for Digital, Culture, Media and Sport, the Home Office and Government Communications Headquarters have collaborated on a recent child safety research project - the Verification of Children Online - that responds to the challenge of platforms knowing which of their users are children. The project engaged with parents and children, industry, regulators and online safety professionals to consider the technical, commercial, legal and behavioural factors that would enable companies to recognise and better protect their child users. A key success of the project was a technical trial run during phase two. The trial successfully demonstrated that age assurance solutions could be run at scale in a way that was simple for users and protects the privacy of their personal data.

- The Verification of Children Online (VoCO) Phase 2 Report was published in November 2020.[35]

Codes of practice

White Paper: *The White Paper stated that the independent regulator would set out how companies could fulfil the duty of care in codes of practice.*

Consultation responses and stakeholder engagement: *Some respondents argued that too many codes of practice would cause confusion, duplication, and potentially, an over reliance on removal of content by risk averse companies.*

Final policy position: *There will not be a code of practice for each category of harmful content. The codes of practice will focus on systems, processes and governance that in-scope companies need to put in place to uphold their regulatory responsibilities. The regulator will decide which codes to produce, with the exception of the codes on child sexual exploitation and abuse and preventing terrorist use of the internet.*

The government will set out high level objectives for the codes of practice with the regulator ensuring that its codes of practice meet these objectives during drafting. Ofcom will consult with relevant parties during the drafting of the codes before sending the final draft to the Secretary of State for Digital, Culture, Media and Sport and the Home Secretary. Ministers will have the power to reject a draft code and require the regulator to make modifications for reasons relating to government policy.

Parliament will also have the opportunity to debate and vote on the objectives and the completed codes will be laid in Parliament.

Due to the seriousness of the harms, and to bridge the gap until the regulator is operational, the government has published interim codes of practice on how to tackle online terrorist and child sexual exploitation and abuse content and activity.

2.48 Ofcom will have a duty to issue statutory codes of practice that set out the steps companies can take to fulfil the duty of care. The codes of practice will focus on systems, processes and governance that in-scope companies need to put in place to uphold their

[35] 'VoCO (Verification of Children Online) Phase 2 report' GCHQ, DCMS, Home Office and ACE, November 2020 (last viewed in November 2020)

regulatory responsibilities. Companies may take alternative steps to those set out in the codes of practice, provided they can demonstrate to Ofcom that those steps are as effective as or exceed the standards set out in the codes.

2.49 Given the range of services in scope of the regulatory framework, some of the steps may not be applicable to every company; conversely the codes will not cover every conceivable risk or emerging technology. If there is no code of practice which covers a particular emerging technology, companies will still need to be compliant with the overarching duty of care. This can be achieved by in-scope companies assessing and responding to the risk associated with those emerging technologies.

2.50 Ofcom will be required to consult with a range of stakeholders when developing codes of practice. This will be critical to ensure that codes take into account existing expertise and best practice regarding how to effectively tackle the range of harmful content and activity in scope of regulation.

Interim codes of practice

2.51 The White Paper committed the government to work with law enforcement agencies and other relevant bodies to produce interim codes of practice on terrorism and child sexual exploitation and abuse due to the serious nature of these harms. The interim codes are voluntary and are intended to bridge the gap until the regulator is operational and ready to produce its own statutory codes on terrorism and child sexual exploitation and abuse, building on the work of the interim codes. The government will work with industry stakeholders to review the implementation of the interim codes so that lessons can be learned and shared with Ofcom, to inform the development of their substantive codes.

2.52 The interim codes are published alongside this response. The government has undertaken an extensive period of engagement across wider government, industry, international partners and civil society, to ensure the measures set out are proportionate but robust enough to tackle these most serious and illegal online harms. As the government is proposing that the interim codes of practice are adopted by all companies in scope, this will include small and medium-sized enterprises. To reduce the burden on businesses and ensure consistency across the industry, the interim codes set out detailed aims and examples of best practice on how to implement each principle.

2.53 The child sexual exploitation and abuse interim code of practice builds on the Voluntary Principles to Counter Online Child Sexual Exploitation and Abuse, that were developed by the UK, US, Canadian, Australian and New Zealand governments, following consultation with tech companies and Non Governmental Organisations.[36]

[36] 'Voluntary Principles to Counter Online Child Sexual Exploitation and Abuse' Five Country Ministerial (last viewed in November 2020)

Using technology to identify illegal child sexual exploitation and abuse content and activity

White Paper: *The White Paper set out that some private channels would be in scope of the online harms regime, however companies would not be required to scan or monitor for illegal content on these services, reflecting the importance of privacy.*

Consultation responses and stakeholder engagement: *Some consultation respondents including industry and civil liberty groups argued that private communications should either fall out of scope or be subject to very limited requirements, to protect user privacy. By contrast, some online safety organisations and children's charities argued private communications should be in scope because there is a high risk of harmful activity - such as child grooming - on private channels.*

Final policy position: *The regulatory framework will apply to both public communication channels and services where users expect a greater degree of privacy. The regulator will set out how companies can fulfil their duty of care in codes of practice, including what measures are likely to be appropriate in the context of private communications. Companies in scope will need to consider the impact on users' privacy and ensure users understand how company systems and processes affect user privacy. The scale, severity and complexity of child sexual exploitation and abuse is particularly concerning, with private channels being exploited by offenders. In light of this, the regulator will have the power to require companies to use automated technology that is highly accurate to identify illegal child sexual exploitation and abuse activity or content on their services. Recognising the importance of protecting users' privacy, the government will ensure this will be used only where there are no alternative measures that are capable of achieving the same aim of reducing harm and subject to stringent legal safeguards to protect users' rights.*

Child sexual exploitation and abuse on private channels

2.54 The government and many stakeholders are particularly concerned about the extent of child sexual exploitation and abuse occurring on some private channels, where offenders believe their illegal activity is less likely to be detected. This could include sharing child sexual abuse material with other offenders, grooming children for sexual purposes, or the livestreaming of abuse.

2.55 These actions cause severe harm and there is clear evidence they are occurring. For example, 12 million of the 18.4 million worldwide child sexual exploitation and abuse reports made by Facebook in 2019 were for content shared on private channels.[37] The identification of this material has real world impact, allowing law enforcement to arrest offenders and safeguard children who would otherwise have been at risk of, or subject to, ongoing abuse. It also protects victims, who can continue to be traumatised long after their abuse by the knowledge that offenders continue to trade in and enjoy the images of their abuse.

[37] 'Tech Companies Detect a Surge in Online Videos of Child Sexual Abuse' The New York Times, February 2020 (last viewed in November 2020)

2.56 Technology that can identify known illegal content accurately and at scale has been in use for many years, supplied by a number of non governmental organisations and web hosting services, and new technology continues to evolve. Some companies already take action against child sexual exploitation and abuse on private channels, but others do not. For example, some companies use PhotoDNA (see Box 14) to identify child sexual abuse material and the gaming sector commonly uses technology to identify harmful activity in communication between users, particularly where a game is aimed at younger children. According to Facebook's community standards enforcement report, in 2019, they actioned 37.4 million pieces of content that violated their child nudity or child sexual exploitation policy. More than 99% of this came to light as a result of the company's proactive efforts (such as use of technology).[38]

2.57 The government has set out in the interim code of practice for online child sexual exploitation and abuse that companies should consider voluntarily using automated technology to identify child sexual exploitation and abuse. The government will continue to support companies using technology to identify online child sexual exploitation and abuse on a voluntary basis once online harms legislation is in force. The government does not intend that restrictions placed on the regulator's power to require a company to use technology should limit companies that choose to go further in taking action.

2.58 Given the serious risk of harm to children, the regulator must have appropriate powers to compel companies to take the most effective action to tackle illegal child sexual exploitation and abuse content and activity on their services, including private communications, subject to stringent legal safeguards.

2.59 Therefore, the regulator will have the express power, where alternative measures cannot effectively address child sexual exploitation and abuse (see 2.58), to require a company to use automated technology that is highly accurate to identify only illegal child sexual exploitation and abuse content or activity on their service. The power is more likely to be considered proportionate on public platforms than on private services. The regulator can take enforcement action if this requirement is not met.

2.60 Robust safeguards will be included in the online harms legislation to govern when the regulator can require the use of automated technology. The regulator will only be able to require the use of tools that are highly accurate in identifying only illegal content, minimising the inadvertent flagging of legal content ('false positives') for human review. The regulator will advise the government on the accuracy of tools and make operational decisions regarding whether or not a specific company should be required to use them. However, before the regulator can use the power it will need to seek approval from Ministers on the basis that sufficiently accurate tools exist. The government assesses that currently, sufficiently accurate tools exist for identifying illegal child sexual exploitation and abuse material that has previously been assessed as being illegal.

2.61 In addition, in order to inform debate around the use of automated technology, the regulator will have to report annually to the Home Secretary and lay a report before Parliament

[38] 'Child Nudity and Sexual Exploitation of Children - transparency' Facebook, 2020 (last viewed in November 2020)

on the use of the power, including the effectiveness and accuracy of the available tools, and any other factors relevant to their suitability for use (for example affordability, availability, and effectiveness).

2.62 In addition to ensuring the accuracy of tools, before requiring a company to use technology to identify child sexual exploitation and abuse, the regulator would need to:

- have gathered evidence which it assesses as demonstrating persistent and prevalent child sexual exploitation and abuse on the service, which the company has failed to address.
- be satisfied that no alternative, less intrusive approaches are available to address the problem and the requirement is proportionate.
- issue a public notice of the regulator's intention to require a company to use automated technology to identify child sexual activity and exploitation, to ensure that users are fully informed.

2.63 In exercising this power, the regulator will balance users' rights to privacy and freedom of expression with the rights of children to be protected from sexual exploitation and abuse.

Box 14: Example of automated technology: PhotoDNA

- One of the technologies commonly used today to identify child sexual abuse material is PhotoDNA, or 'hash matching'. This converts images into a numerical code (or hash) that can be compared against the codes for known images of child sexual abuse.

- This technology is only capable of assessing whether an image is child sexual abuse, and makes no other inferences about the image or user's communication. When a match is detected, the image can be reviewed, blocked, taken down or reported by the company.

- The false positive rate is estimated to be between one in two billion and one in ten billion, protecting the privacy of legitimate users whilst ensuring no safe space for child sexual abuse offenders to operate.[39]

- A range of non governmental organisations and web hosting providers make this technology and the hash data sets available to companies looking to protect their service from abuse.

Using technology to identify terrorist content and activity on public services

White Paper: *The White Paper set out that the regulator would not compel companies to undertake general monitoring on their online services, as this would place a disproportionate burden on companies and raise concerns about freedom of expression and user privacy.*

[39] https://www.itnews.com.au/news/facebook-deploys-photodna-to-scan-for-child-abuse-material-258301#:~:text=According%20to%20Dartmouth%20computer%20scientist%20Hany%20Farid%2C%20PhotoDNA,false%20positive%20rate%20of%20between%20one%20in%20

Instead, the new regulatory framework would increase the responsibility of online services in a way that is compatible with the European Union's e-Commerce Directive, which limits their liability for illegal content until they have knowledge of its existence, and have failed to remove it from their services in good time. However, it noted the strong case for mandating specific monitoring for tightly defined categories of illegal content where there is a threat to national security or the physical safety of children.

Consultation responses and stakeholder engagement: *Industry welcomed the commitment to maintaining existing intermediary liability provisions set out in the e-Commerce Directive, including the prohibition on general monitoring.*

Final policy position:
Many companies already use technology to identify and remove illegal terrorist content from their services. The regulator will also be given an additional express power in legislation, to require a company to use that technology to identify and remove illegal terrorist content from their public services where this is the only effective, proportionate and necessary action available, and the regulator is confident that the tools available are highly accurate at identifying only illegal content to minimise the need for human review of legal content.

Companies' liability for specific pieces of content will remain unchanged. Once a company is aware of illegal content on their service, it will still be required to take this down quickly otherwise it could become liable for that content. Where technology is used to identify the tightly defined categories of illegal content set out above, companies' will therefore need to remove it to avoid incurring liability. The technology used will be highly accurate and therefore unlikely to identify illegal content that does not constitute an offence relating to terrorism. This applies equally to the requirements relating to child sexual exploitation and abuse, set out above.

2.64 The White Paper set out the reasonable steps that companies should take in advance of legislation to prevent new and known terrorist content and activity on their services. This included the proactive use of automated technology, where appropriate, to identify, flag, block or remove illegal content and activity.

2.65 The government has set out in the interim code of practice for online terrorist content and activity that companies should consider voluntarily using automated technology to identify and remove terrorist content and activity from their public services. The government will continue to support companies using technology to identify online terrorist content and activity on a voluntary basis once online harms legislation is in force.

2.66 The regulator will also be given an additional express power in legislation to require a company to use such automated technology to identify and remove illegal terrorist content from their public channels, where this is the only effective, proportionate and necessary action available. The regulator can take enforcement action if this requirement is not met.

2.67 This power will be used only if (i) the technology is highly accurate in identifying illegal terrorist content (ii) there is evidence of persistent and prevalent illegal terrorist activity on public channels of a service and (iii) other measures could not be equally effective. As with the rest of the online harms framework, any requirements resulting from this power must be proportionate.

2.68 Automated technologies are already employed by some companies on a voluntary basis, as part of their own efforts to tackle terrorist content and activity on their services. However, this is not done widely or consistently.

2.69 Companies also rely on user reports or referrals from law enforcement to alert them to content already on their services, so that they can remove it (if illegal or breaching their terms and conditions). These reports also help fine-tune their automated tools. However, reactive measures such as those set out above cannot by themselves adequately tackle the speed and scale with which terrorist content online is often disseminated. Referrals from the Counter Terrorism Internet Referral Unit successfully led to over 310,000 individual pieces of terrorist content being removed by companies between 2010 and the end of 2018,[40] but transparency reports indicate that this is just a fraction of what companies can action proactively on their own services. For example, Facebook reported that between April and June 2020, 8.7 million pieces of terrorist content were actioned, 99.6% of which were found and flagged by Facebook before users reported it.[41]

2.70 The safeguards built into the regulation, detailed in paragraph 2.62, will ensure the approach to terrorist content and activity on public services is proportionate – balancing taking action against illegal terrorist content and activity in the interests of protecting national security and upholding users' rights online.

Data retention and reporting to law enforcement

White Paper: *The White Paper stated that the regulator would provide specific guidance in its code of practice on the content companies should preserve following removal and for how long. It also set out that the regulator would provide guidance on when companies should proactively alert law enforcement and other relevant government agencies about specific illegal content.*

Consultation responses and stakeholder engagement: *Stakeholders, including the National Crime Agency and National Centre for Missing and Exploited Children, argued that there should be new, mandatory reporting requirements for child exploitation and sexual abuse content to increase reporting and standardise the approach. In their view, this will improve the ability of law enforcement to tackle child sexual exploitation and abuse offenders and safeguard victims in the UK and elsewhere.*

Final policy position: *The government is minded to introduce a requirement for companies to report child sexual exploitation and abuse identified on their services, with these reports being made to a designated body. A requirement to retain child sexual exploitation and*

[40] 'Together, We're Tackling Online Terrorism' Counter Terrorism Policing, December 2018 (last viewed in November 2020)
[41] 'Dangerous Organizations: Terrorism and Organized Hate - transparency' Facebook, 2020 (last viewed in November 2020)

abuse data will not be introduced through this legislation. However the government is considering introducing this through alternative legislation.

With regards to terrorist content and activity, the government expects companies to report to law enforcement where they consider there is a threat to life or risk of imminent attack. The legislation will not introduce a requirement for companies to retain this data.

2.71 The White Paper indicated that the regulator would be expected to set out in the terrorist and child sexual exploitation and abuse codes of practice the reasonable steps that companies could take in relation to retaining data and reporting these types of content. The regulator would include guidance on how long companies should retain data for and the circumstances in which content should be reported to law enforcement and other agencies.

2.72 Following the White Paper consultation and further engagement with law enforcement and other agencies, the government is minded to introduce a mandatory requirement on companies to report child sexual exploitation and abuse identified on their services. Further work is being undertaken to explore a suitable body to receive these reports and to ensure this system does not duplicate companies' existing reporting obligations. This would be a standalone legislative requirement, rather than part of the duty of care. This approach reflects the seriousness of this crime and seeks to ensure that companies provide high quality reports with the information law enforcement need to identify offenders and safeguard victims.

2.73 Companies will be encouraged to retain child sexual exploitation and abuse data for law enforcement purposes. The online harms legislation will not introduce a requirement to retain this data but the government is considering introducing this requirement within alternative legislation.

2.74 The government expects companies to report terrorist content and activity on their services to law enforcement where they consider there is a threat to life or risk of imminent attack. The government will work with the regulator to ensure that it encourages this and provides companies with clear guidance on how this could best be done and information on where to report to. The online harms legislation will not introduce a legal requirement for companies to report and retain this data.

Disinformation and misinformation

White Paper: *The White Paper did not set out a definitive position on how disinformation and misinformation would be addressed under the regulatory framework. Disinformation was included in an indicative list of harmful content or activity that would be within scope of the legislation, because it can be harmful to both individuals and to society.*

Consultation responses and stakeholder engagement: *A range of stakeholders, including civil society organisations, raised concerns about including disinformation and misinformation in scope of the regulation because of the impact this might have on freedom of expression. Many stakeholders are concerned about the threat that disinformation and* *misinformation poses to individual users, as well as its potential broader impact on public safety, national security and community cohesion.*

Final policy position: *Companies will need to address disinformation and misinformation that poses a reasonably foreseeable risk of significant harm to individuals (e.g. relating to public health).*

The legislation will also introduce additional provisions targeted at building understanding and driving action to tackle disinformation and misinformation. For example, establishing an expert working group on disinformation and misinformation, measures to improve transparency about how companies deal with disinformation and building on Ofcom's existing duties to promote media literacy.

Where disinformation and misinformation presents a significant threat to public safety, public health or national security, the regulator will have the power to act.

2.75 The White Paper set out the dangers of online disinformation and misinformation to both individuals and society. Disinformation is the deliberate creation and dissemination of false and/or manipulated information that is intended to deceive and mislead audiences, either for the purposes of causing harm, or for political, personal or financial gain. Misinformation refers to inadvertently spreading false information.

2.76 COVID-19 has brought these dangers into sharp focus. Ofcom data suggests that in week one of the UK lockdown, nearly 50% of respondents reported seeing information they thought to be false or misleading about the pandemic, with this figure at almost 60% for 18-34 year old respondents.[42] While Ofcom has recorded a gradual decrease in self-reported exposure to narratives considered false or misleading, navigating a COVID-19 online environment can be challenging and at times, confusing for many people in the UK.[43]

2.77 The government is taking a range of steps to tackle disinformation and misinformation online. In response to the pandemic, the government stood up the Department for Digital, Culture, Media and Sport-led cross-Whitehall Counter Disinformation Unit, to provide a comprehensive picture of the extent, scope and the reach of disinformation and misinformation, and to work with partners to ensure appropriate action is taken. Since standing up, the Unit has observed a range of false narratives, some of which have caused significant harm to individuals and society. Examples include conspiracy theories inaccurately linking COVID-19 with 5G technologies, health misinformation promoting a range of junk cures, and stories using outdated footage to suggest certain groups were breaking social distancing.

2.78 As the pandemic has progressed, the Unit has also seen other narratives gain traction, particularly those which seek to undermine efforts to produce a COVID-19 vaccine. Anti-

[42] 'Covid-19 new and information: summary of views about misinformation' Ofcom, June 2020 (last viewed in November 2020)

[43] Concepts and definitions of misinformation can be partial and subjective, and often depend upon the respondent's own sets of beliefs and affiliations.The survey data relies on self-reported exposure and is, therefore, unlikely to represent the true proportion of the population exposed to COVID-19 misinformation

vaccination disinformation and misinformation has the potential to cause significant harm to individuals. Given the pace at which such narratives can develop on social media, combined with established movements against inoculation, reducing the risk of such harm remains a key priority. The Department for Digital Culture, Media and Sport is working with cross-Whitehall partners, particularly the Department for Health and Social Care, and social media services, to mitigate and tackle the risk of anti-vaccination false information.

2.79 Coupled with these efforts, the government has continued to build audience resilience to disinformation and misinformation, enabling people to critically assess, appraise and challenge information online. Through the 'Don't Feed the Beast' campaign and SHARE checklist, UK users have been given five easy steps to identify false content, encouraging them to consider information they share online. The forthcoming online media literacy strategy (see below and Part 5 for more information on Media Literacy) will set out more action to improve and strengthen audience resilience. Under the Cabinet Office led Defending Democracy programme, the government is also taking further steps to strengthen the integrity of UK elections and promote fact-based and open discourse. This includes responding to recommendations on press sustainability made in the Cairncross review (see Box 22), and the introduction of a digital imprints regime.

Disinformation and misinformation under the new regulatory framework

2.80 Legislation has an important part to play in tackling this harm. The White Paper included disinformation in the indicative list of harmful content or activity that would be within scope of the legislation, because it can be harmful to both individuals and to society.

2.81 As set out in paragraph 2.2, the duty of care will apply to content or activity which could cause significant physical or psychological harm to an individual, including disinformation and misinformation. Where disinformation is unlikely to cause this type of harm it will not fall in scope of regulation. Ofcom should not be involved in decisions relating to political opinions or campaigning, shared by domestic actors within the law.

2.82 Under our proposals, disinformation and misinformation that could cause significant harm to an individual will be within scope of the duty of care. The vast majority of disinformation and misinformation is legal, while potentially harmful. As an example, this would include content which suggests that users should go against established medical advice, such as avoiding vaccinations. There may also be some cases where disinformation is illegal and could cause significant harm to individuals - for example, disinformation which directly incited harm against individuals. In these cases, companies would be expected to remove such content.

2.83 Some types of legal but harmful disinformation and misinformation are likely to be proposed in secondary legislation as categories of priority harm that companies must address in their terms and conditions. Companies must also risk assess for categories of emerging harm. As with other legal but harmful content, companies providing Category 1 services will need to make clear what is acceptable on their services for such content in their terms and conditions and will be required to enforce this. Companies whose services are likely to be accessed by children will also need to take steps to protect children from disinformation and misinformation which could be harmful to them.

2.84 As the pandemic has demonstrated, there may be instances when urgent action is required to address disinformation and misinformation during emergency situations. Where disinformation and misinformation presents a significant threat to public safety, public health or national security, the regulator will have the power to act. In such situations, Ofcom will be able to take steps to build users' awareness and resilience to disinformation and misinformation, or require companies to report on steps they are taking in light of such a situation.

2.85 To ensure the future regulatory framework is well equipped to deal with the longer-term challenges presented by disinformation and misinformation, the regulator will be required to establish an expert working group on disinformation and misinformation. The working group will aim to build consensus and technical knowledge on how to tackle disinformation and misinformation. This working group will include a range of stakeholders such as rights groups, academics and companies.

2.86 The regulatory framework will also help build an understanding of what companies are doing in relation to disinformation and misinformation through transparency reporting requirements. As set out in the transparency section, the regulator will have the power to require certain companies to publish annual transparency reports, setting out the extent and response to this harm. As part of this, companies could be required, where relevant, to report on processes and systems in place to respond to disinformation and misinformation.

2.87 The regulatory framework will build on Ofcom's existing duties to promote media literacy. This will help increase user awareness of, and resilience to, disinformation and misinformation online (for more information on Media Literacy, see Part 5).

2.88 The government has also committed to publishing a safety by design framework (see Part 5). This will set out best practice and specific measures that companies can take to address the risk of harm on their services. This will include design measures to address the risk of misinformation and disinformation spreading on services, and empower users to engage critically with information online.

Part 3: The regulator

Summary

Consultation questions covered in Part 3:

- ❖ *What role should Parliament play in scrutinising the work of the regulator, including the development of codes of practice?*
- ❖ *Should an online harms regulator be: (i) a new public body, or (ii) an existing public body? If your answer to question 10 is (ii), which body or bodies should it be?*
- ❖ *A new or existing regulator is intended to be cost neutral: on what basis should any funding contributions from industry be determined?*

- The government can now confirm that Ofcom will be named as the online harms regulator in legislation. Ofcom has a strong strategic fit for this role, and relevant organisational experience as a robust independent regulator. Empowering an existing regulatory body will help the timely introduction of the online harms regime by allowing Ofcom to begin preparations now to take on the role.

- Ofcom will raise the required income to cover the costs of the regime from industry.

- The regulator will be accountable to Parliament. Ofcom as the regulator will lay its annual report and accounts before Parliament and be subject to Select Committee scrutiny. The annual report will give details about how it has discharged its functions in relation to online harms.

Body (new vs. existing) and identity of regulator

White Paper: *The White Paper stated that the online harms regime will be overseen and regulated by an independent regulator. It also explained that the government would consider whether a broader restructuring of the regulatory landscape would reduce the risk of duplication and minimise burdens on business.*

Consultation responses and stakeholder engagement: *The responses emphasised the need for there to be consistency between existing and new regulatory regimes, and for the regulator to be equipped to function effectively. Views on the identity of the regulator were balanced, highlighting the benefits and risks of a new body versus an existing one.*

Final policy position: *In February 2020, the government announced that it was minded to give Ofcom the role of the independent online harms regulator. The government can now confirm that Ofcom will be named as the regulator in legislation. Empowering an existing regulatory body will help the timely introduction of the online harms regime, by allowing Ofcom to begin preparations now to take on the role.*

3.1 To inform the set up of the independent online harms regulator, the consultation asked questions about its identity, funding model and accountability to Parliament. The government has examined a range of options including creating a new body or appointing an existing regulator. These options were assessed against a number of key criteria, including effectiveness, efficiency and strategic coherence, and were informed by feedback from the consultation response.

3.2 In February 2020, the government announced that it was minded to give Ofcom the role of the independent online harms regulator. Ministers have now decided to confirm the appointment of Ofcom to this role, subject to the passage of legislation. This preference was based on its organisational experience, robustness, and experience of delivering whilst holding challenging, high-profile remits across a range of sectors. Ofcom also offered a strong strategic fit given its role regulating activities increasingly related to online harms, and their new responsibilities in relation to regulating UK-established video sharing platforms.

3.3 Ofcom was established by the Office of Communications Act 2002 from the convergence of five existing communications regulators covering broadcasting and telecommunications, and received its full authority from the Communications Act 2003. Since then, it has had other duties added to its remit, including postal services in 2011 and the BBC in 2017. The technological revolution of traditional communication industries has meant that digital and online services have increasingly become part of Ofcom's existing remit. It is therefore well placed to play a similar role for online harms.

3.4 Whilst meeting the challenge of online harms requires new ideas, it is also vital that the government utilises the experience, expertise and infrastructure of the UK's existing world class regulators. Ofcom has an existing network of relationships in the tech sector, experience of dealing with a high volume of small businesses, and a research-led, risk-based approach to regulation. This provides a strong foundation for taking on the online harms regime.

3.5 Earlier this year, the government announced Ofcom as the national regulator for UK-established video sharing platforms under the Audiovisual Media Services Regulations. These came into force on 1 November 2020. The regulations introduce new requirements for UK-established video sharing platforms to protect users from harmful content. In the longer term, the government intends for the regulation of UK-established video sharing platforms to be part of the online harms regime. This alignment between the two regimes offers the opportunity for early engagement with stakeholders and for testing regulatory processes ahead of the online harms legislation coming into force. Ofcom's increasing role in regulating activities relating to online harms further emphasises its strong strategic fit to be the independent online harms regulator.

Box 15: Audiovisual Media Services Regulations 2020

- The UK's Audiovisual Media Services Regulations 2020 place requirements on UK-established video sharing platforms to protect their users from certain types of harm.

- The regulations include a requirement for UK-established video sharing platforms to take appropriate measures to protect children from harmful content, and to protect the general public from incitement to hatred and violence and from criminal content. They also include requirements relating to standards around advertising. The statutory framework was introduced into legislation in Autumn 2020 and came into force from 1 November 2020. Ofcom is actively engaging with providers of video sharing platforms, and will be developing and publicly consulting on regulatory guidance for platforms in the coming months.

- The regulations share broadly similar objectives to the online harms regime. The government's preference is for the requirements on UK-established video sharing platforms to transition to, and be superseded by, the online harms regulatory framework, once the latter comes into force. Under the online harms regulatory framework, UK-established video sharing platforms will continue to have systems and processes in place to protect users.

- The requirements on UK-established video sharing platforms in relation to audiovisual commercial communications under the Audiovisual Media Services Regulations 2020, will also be repealed and will not be encompassed in the online harms regime. This is because the Advertising Standards Authority's self-regulatory rules already apply equivalent standards for advertising as those in the regulations. The Advertising Standards Authority's rules require all online advertisers to adhere to specific advertising standards. Even after the requirements of the revised regulations have been subsumed by the online harms regime, the Advertising Standards Authority's rules will continue to apply to all online advertisers.

- In tandem with the online harms work, the Department for Digital, Culture, Media and Sport is currently engaged in a programme of work related to online advertising which, amongst other areas of focus, is looking at ensuring that advertising regulation answers the needs of the changing advertising marketplace. Further details on this are set out in Box 3.

3.6 Ofcom has a strong track record of engagement. Its annual report details how it seeks to understand consumers' and citizens' interests and behaviours, and how it engages with industry and government. Successful delivery of the online harms regime will require being able to clearly communicate the purpose and reach of the regulatory framework and the regulator's role, as well as listening to others. The regulator will be required to take a consultative approach, including on the production of codes of practice. The legislation will introduce a super-complaints function and user advocacy mechanisms (see Part 4). Users will also be able to report their concerns to the regulator, however, the regulator will not investigate or arbitrate on individual cases. This would conflict with the principle of a systems and processes approach, and could overwhelm Ofcom, given the likely volume of complaints. Instead, receiving user complaints will be an essential part of Ofcom's activity to ensure the regulator is actively listening to users' experiences and addressing their concerns.

3.7 The government is confident that Ofcom is the right organisation to deliver the online harms regulatory regime, subject to final parliamentary approvals. It is a well-established regulator with a strong reputation internationally and a proven track record of taking evidence-based decisions which balance robust consumer protection with the need to ensure the regulatory environment is conducive to economic growth and innovation. It is sensitive to the need not to impose unnecessary burdens on businesses, and is well versed in best regulatory practice of being transparent, accountable, proportionate and consistent, taking action only when it is needed. Ofcom also has extensive experience of acting in the interests of all UK citizens and consumers, with offices in each of the nations.

3.8 Ofcom's extensive experience of regulating the communications sector in the UK means that this new role is a logical extension of its existing remit. Crucially, Ofcom's experience in broadcasting regulation in particular means that it is well practised in understanding and making judgements about the balance between protection from harm and upholding freedom of expression. Overall, Ofcom is a strong strategic fit for the role.

3.9 There are advantages of working with an existing regulatory body, when compared to creating a new body. Whilst the government remains responsible for the overall policy and for creating the legislation, it is actively engaging with Ofcom to seek the benefit of its regulatory expertise and experience in understanding how the regulatory framework will work in practice. The government expects this process to continue as the legislation is prepared for introduction in Parliament.

3.10 Ahead of its role being confirmed in legislation, Ofcom must seize the opportunity to prepare organisationally and to build on the opportunities provided by its current responsibilities as the national regulator for UK-established video sharing platforms. Ofcom will be able to further its engagement with companies whose services will be in scope of the online harms regime. It will be able to set out the expectations on companies and ensure a fuller understanding of what compliance will entail ahead of the duty of care coming into force.

Governance, capabilities and infrastructure

White Paper: *The White Paper stated that the regulator will be an independent body and that the government will take steps to ensure that the regulator can command public confidence in its independence, impartiality, capability and effectiveness.*

Consultation responses and stakeholder engagement: *Most respondents to the consultation viewed an independent and empowered regulator as critical to delivering the regime. There were no particular views on the regulator's governance arrangements.*

Final policy position: *The online harms legislation will maintain Ofcom's organisational independence and governance arrangements, and clearly define the respective roles of government and the regulator.*

3.11 The importance of regulators being independent from undue influence - from government, other political sources, regulated services and organisations with an interest in the regulated area - is an important element of effective regulation. The White Paper stated that the regulator will be an independent body and that the government will take steps to ensure that the regulator can command public confidence in its independence, impartiality, capability and effectiveness. This will be important for online harms regulation, particularly to manage concerns about protecting freedom of expression.

3.12 Ofcom's founding legislation already provides it with a high degree of independence as it is operationally independent from government, giving it the statutory provisions to manage its own affairs. The government will set a clear scope for the regime and remit of the regulator in legislation. It will give the regulator a high level of independence over fulfilling its duties, and delivering the functions set out in Box 16.

3.13 In some areas, such as the production of codes of practice and the threshold for companies in scope to pay the annual fee, the government will maintain levers to ensure the policy intent of the regulatory framework is maintained. The government will introduce a power to allow the Secretary of State for Digital, Culture, Media and Sport to issue guidance to the regulator, with clearly defined scope and use. This will enable the government to set out further detail on regulatory processes, but will not stray into operational matters or seek to fetter Ofcom's independence in how it operates the regime. The final version of this guidance will be subject to parliamentary approval.

3.14 There will also be an option for the Secretary of State for Digital, Culture, Media and Sport to issue a Statement of Strategic Priorities in relation to the regulatory framework. This power will be similar to the existing powers the Secretary of State has in relation to telecommunications, the management of the radio spectrum, and postal services. This will allow the government to be clear on the overall strategic direction for tackling online harms and to respond at a high level to future changes. The Statement of Strategic Priorities will require external consultation (including with Ofcom) and approval by Parliament. It is not intended for this to be in place, or be needed, at the outset of the regime. Its main aim will be to cater for changes in the digital and regulatory landscape.

3.15 The Secretary of State for Digital, Culture, Media and Sport appoints the non-executive members of the Ofcom Board including the chair, and will work with Ofcom to ensure that the Board has the necessary skills and expertise as it takes on these new responsibilities.

3.16 In addition to this, it is vital that Ofcom has the right skills and expertise to discharge the responsibilities it will have effectively. For example, Ofcom has recognised that it will need more expertise in technology, especially emerging tech, the use of Artificial Intelligence and how this is driving commercial and consumer change. This will be increasingly relevant to all of Ofcom's sectors. It will be needed regardless of Ofcom's online harms role, albeit at a different scale, and ultimately by all regulators overseeing activities and sectors that have an increasingly digital dimension. Ofcom is in the process of building these capabilities and has created a new Emerging Technology directorate and data science team as part of its efforts to do so. Its confirmation as regulator will enable it to drive forward its plans to recruit the relevant experts in these areas, including a new Chief Technology Officer.

Accountability to Parliament

White Paper: *The White Paper stated that it will be important to ensure that Parliament is able to scrutinise the regulator's work.*

Consultation responses and stakeholder engagement: *Responses to the consultation question showed strong support for parliamentary oversight. Most stakeholders agreed that Parliament should not interfere with the regulator's independence in drafting codes of practice. Several responses suggested establishing a dedicated body for reviewing codes.*

Final policy position: *The regulator will be accountable to Parliament for its regulatory activities, including specific aspects of the regime beyond primary legislation.*

3.17 The section above sets out ways in which the regulator will be accountable to the government. In addition it will be accountable to Parliament. Ofcom, as the regulator, will lay its annual report and accounts before Parliament and be subject to Select Committee scrutiny. This will include the chair and senior managers appearing before Select Committees as well as pre-appointment scrutiny for the chair by the Department for Digital, Culture, Media and Sport Select Committee. This is in line with Ofcom's current arrangements.

3.18 Parliament will also have a role in approving a number of aspects of the regulatory framework through its scrutiny of secondary legislation. This will include the statutory instruments establishing the objectives set by the Secretary of State for Digital, Culture, Media and Sport for the codes of practice, the codes of practice themselves and the priority categories for harms.

3.19 The Secretary of State for Digital, Culture, Media and Sport will undertake a review of the effectiveness of the regime 2-5 years after entry into force, producing a report which will then be laid in Parliament. Parliament will have an opportunity to debate the findings of the report.

3.20 To ensure regulatory requirements are proportionate, Ofcom will be required to conduct and publish impact assessments for proposals which will affect businesses in scope of the legislation. This will include codes of practice (detailed in Part 4), but may also include other policy areas such as enforcement, information gathering, transparency, super-complaints, media literacy or funding. The regulator will have a specific duty to assess the impact of its proposals on small and micro businesses, to ensure undue regulatory burdens are not placed on them. It will be required to consult on impact assessments, to ensure it is gathering the best available evidence and to provide transparency. Ofcom will also be required to report on the impact assessments it has undertaken in annual reports to Parliament.

Regulator funding model

White Paper: *The White Paper and the initial government response both outlined that the* *regulator will be funded by industry in the medium term. The government indicated it would consider a range of options to fund regulator activity, including fees, charges and levies on services in scope.*

Consultation responses and stakeholder engagement: *There was broad agreement amongst stakeholders and respondents to the consultation, that funding should primarily be from industry. However, it was felt that the model should be proportionate and practical for example, by minimising unnecessary costs on smaller businesses and ensuring efficient collection of contributions from companies based overseas.*

Final policy position: *Ofcom will be given powers to raise the required income to cover the costs of running the online harms regime from industry.*

3.21 Ofcom will be able to raise the required income to cover the costs of running the online harms regime through industry fees. Ofcom will also have the power to require a company to undertake, and pay for, a skilled person report.

Notification and the Annual Fee

3.22 Companies above a threshold based on global annual revenue will be required to notify the regulator and pay an annual fee. Companies below the threshold will not be required to notify the regulator or pay a fee. The threshold will be set by Ofcom, based on consultation with industry, and will be signed off by Ministers. Companies in scope which fall below the threshold will still have to comply with all their other regulatory responsibilities. The regulator will, in consultation with industry, prescribe the details of the notification process, including the information required from the company at the point of notification.

3.23 The total amount of fees to be charged to industry will be in proportion to the costs incurred by the regulator in operating the online harms regime. The fees to be paid by individual companies will be tiered. The intention is that the regulator will calculate the fees based on two metrics: a primary metric of global annual revenue; and a secondary optional metric based on company activity. The details of the second metric will be determined by the regulator and could be calculated using criteria such as the presence of specific functions on a service. The metrics used to calculate the fees will meet the strict criteria of proportionality, affordability and objectivity.

3.24 The government will give consideration to how all costs will be managed within this funding regime, including the litigation costs of the regulator and will work to ensure that the regulator remains cost neutral to the taxpayer.

3.25 As detailed in the section on Information gathering and Investigation, the regulator will have the power to require a skilled person report on specific issues of concern.

3.26 When the regulator uses this power, the company will always be required to cover the direct costs of the skilled person report. The regulator will consider the use of alternative powers to obtain information it needs, if it determines that paying for the skilled person report could have an adverse financial impact on the company.

Interface with other bodies

White Paper: *The White Paper stated that the government and regulator will need to work closely with a number of other organisations, both domestic and international, to ensure the successful implementation of the online harms regime. For example, industry bodies, other regulators, law enforcement and overseas bodies.*

Consultation responses and stakeholder engagement: *The consultation responses showed strong support for coordination and cooperation across regulators. The emphasis was on UK based regulators as the questions did not reference international engagement.*

Final policy position: *The government will work with Ofcom to ensure that the regulator is able to work effectively with a range of organisations. This will be delivered through a range of means including co-designation powers, memorandums of understanding, forums and networks.*

3.27 The online harms regulator will need to have a large number of relationships with other organisations, including regulators, government bodies, the devolved administrations, public agencies, industry bodies, law enforcement and civil society. Ofcom already has a strong network of relationships with a range of bodies and will continue to cultivate these both at home and internationally. These relationships allow Ofcom to draw on the expertise within other bodies whilst maintaining its own independence.

3.28 These relationships, variously underpinned by memorandums of understanding, forums and networks, will support the regulator in understanding the prevalence of and impact of online harms and the effectiveness of companies' responses. Relationships with civil society will be critical to ensure the regulator understands the needs of different user groups. In turn, this will help ensure that the regulatory framework adequately keeps pace with the online threat. Furthermore, Ofcom will play a critical role in enforcement across borders, and will rely on its good relationships with its international counterparts to facilitate obtaining information from other jurisdictions, and to achieve a degree of international regulatory alignment.

3.29 Ofcom will have the power to co-designate other bodies to deliver aspects of the regulatory framework to make use of the significant expertise that sits outside Ofcom. The government will work with Ofcom to understand where this may be effective and beneficial to delivering the regulatory framework.

3.30 It will also be important to ensure that the regulator and law enforcement are able to share information with each other as appropriate to support the delivery of their functions. Ofcom will need to build strong working relationships with law enforcement and other agencies in order to, among other things, develop their understanding of, and take effective action against, online terrorism and extremism.

3.31 Ofcom has strong existing relationships with other regulators such as the Information Commissioner's Office and the Competition and Markets Authority. On 1st July 2020, the Competition and Markets Authority, Information Commissioner's Office and Ofcom announced a new forum to help ensure online services work well for consumers and businesses in the

UK. The Digital Regulation Cooperation Forum aims to strengthen existing collaboration and coordination between the three regulators by harnessing their collective expertise when data, privacy, competition, communications and content interact. Ofcom will work closely with these and other regulatory bodies to coordinate various aspects of digital regulation including online harms.

3.32 The decision to appoint Ofcom as the regulator for online harms is part of a wider programme of work the government is undertaking to ensure the regulatory landscape for digital technologies is coherent, effective and efficient.

3.33 The government will continue to review and assess the regulatory landscape as new powers across our digital regulation programme are proposed. Where necessary, the government will introduce further reform to ensure our institutions are fully coherent, efficient and effective.

Part 4: Functions of the regulator

Summary

Consultation questions covered in Part 4:

- ❖ *This government has committed to annual transparency reporting. Beyond the measures set out in the White Paper, should the government do more to build a culture of transparency, trust and accountability across industry and, if so, how?*
- ❖ *Should designated bodies be able to bring 'super-complaints' to the regulator in specific and clearly evidenced circumstances? In what circumstances should this happen?*
- ❖ *What, if any, other measures should the government consider for users who wish to raise concerns about specific pieces of harmful content or activity, and/or breaches of the duty of care?*
- ❖ *Should the regulator be empowered to i) disrupt business activities, or ii) undertake ISP blocking, or iii) implement a regime for senior management liability? What, if any, further powers should be available to the regulator?*
- ❖ *Should the regulator have the power to require a company based outside the UK and EEA to appoint a nominated representative in the UK or EEA in certain circumstances?*
- ❖ *In addition to judicial review should there be a statutory mechanism for companies to appeal against a decision of the regulator, as exists in relation to Ofcom under sections 192-196 of the Communications Act 2003? In what circumstances should companies be able to use this statutory mechanism? Should the appeal be decided on the basis of the principles that would be applied on an application for judicial review or on the merits of the case?*
- ❖ *What, if any, advice or support could the regulator provide to businesses, particularly start-ups and small and medium-sized enterprises, to comply with the regulatory framework?*
- ❖ *What further steps could be taken to ensure the regulator will act in a targeted and proportionate manner?*

- The regulator's primary objective will be to improve safety for users of online services. It will undertake regulatory action in line with the principles of the regulatory framework.

- The regulator will set out what companies need to do to fulfil the duty of care, including through codes of practice.

- The regulator will have the power to require certain companies in scope to publish annual transparency reports, which will empower users to make informed decisions about which services they use.

- The regulator will be able to access information about companies' redress mechanisms in the exercise of its statutory functions, and will accept complaints from users as part of its horizon-scanning and supervision activity.

- The regulator will have a super-complaints function, and will accept super-complaints when there is substantial evidence of a systemic issue affecting large numbers of people, or specific groups of people. The regulator will also establish appropriate mechanisms for user advocacy, to ensure users' experiences and concerns are being heard and acted upon.

- The regulator will have a range of robust enforcement powers to tackle non-compliance by in-scope companies providing services to UK users, to ensure the effectiveness of the regime.

Duties on and functions of the regulator

4.1 The regulator will have certain duties and functions under the framework. Its primary duty will be to improve the safety of users of online services (and that of non-users who may be directly affected by others' use of them). Regulatory action should be undertaken in line with the principles of the regulatory framework (see Annex A), which means being realised in a way that:

- is based on the risk of content or activity online harming individuals, where it gives rise to a reasonably foreseeable risk of a significant adverse physical or psychological impact on individuals;
- is reasonable and proportionate to the severity of the potential harm and resources available to companies;
- provides a higher level of protection for children than for adults;
- protects users' rights, including to freedom of expression and privacy online; and safeguards media freedom;
- promotes transparency about and accountability for the incidence of and response to harm;
- supports innovation and reduces the burden on business; and
- is delivered by putting in place appropriate systems and processes.

4.2 Ofcom, as the regulator, will need to apply these principles when it issues codes of practice which will set out steps companies can take to fulfil the duty of care.

4.3 In addition to the above, Ofcom will also need to pay due regard to innovation in the exercise of all of its functions, and it will have a further responsibility to help all companies to understand and fulfil their responsibilities. This will involve providing appropriate support to companies depending on their size and maturity, with greater help for small and medium-sized enterprises. It will also be required to assess the impact of its regulatory activities on business, and in particular small and micro businesses.

4.4 Under the online harms regime Ofcom will have a duty to consider the vulnerability of children and of others whose circumstances appear to Ofcom to put them in need of special protection when performing its duties.

Box 16: Regulator role and functions

The regulator's role and functions will include:

- Setting out what companies need to do to fulfil the duty of care, including through codes of practice.
- Establishing a transparency, trust and accountability framework.
- Requiring all in-scope companies to have effective and accessible mechanisms for users to report concerns and seek redress for alleged harmful content or activity online, infringement of rights, or a company's failure to fulfil its duty of care.
- Assessing and responding to super-complaints.
- Establishing user advocacy mechanisms to understand users' concerns and experiences.
- Taking prompt and effective enforcement action in the event of non-compliance, when it is appropriate and proportionate.
- Providing support to start-ups and small and medium-sized enterprises to help them fulfil their legal obligations in a proportionate and effective manner.
- Promoting education and awareness-raising about online safety to empower users to stay safe online.
- Undertaking and commissioning research to improve our understanding of online harms, their impacts on individuals and society and how they can be tackled.

Codes of practice

4.5 The government will set objectives for the regulator's codes of practice in secondary legislation to provide clarity for the framework. Ofcom will have a duty to consult interested parties on the development of the codes of practice, which is consistent with usual regulatory practice. It will also be required to consult bodies, organisations and interests specified in legislation who have specific knowledge and expertise relating to the policy objectives, or who have a significant interest in the online harms regime. The government will require the regulator to undertake impact assessments for both new codes of practice and for revisions to existing codes. As referenced at paragraph 3.20, this will include a specific requirement to assess the impact of codes of practice on small and micro businesses. This will help to ensure regulatory requirements are proportionate and that they do not place an undue burden on businesses.

4.6 There will not be individual codes of practice for each specific harm; it will be for the regulator to decide which codes of practice to produce. There are some exceptions to this, where codes of practice will need to be more focussed. For example, there will be individual codes of practice on tackling terrorist use of the internet, and on child sexual exploitation and abuse. This reflects the requirement on companies to take particularly robust action on these issues.

4.7 The government will maintain levers to ensure that the policy intent of the framework is upheld and that evidence and expertise from government and law enforcement agencies is reflected in the codes. For example, ministers will be statutory consultees for the codes of practice.

4.8 Ministers will have the power to issue a direction to reject a draft code for reasons relating to government policy. Ministers would, where appropriate, publish the letter of direction to the regulator, which would also set out modifications the regulator must make when revising the code. The power could be used only at the end of the drafting process when the codes are submitted by Ofcom to the Secretary of State for Digital, Culture, Media and Sport and the Home Secretary. Ofcom will be responsible and accountable for all the codes of practice.

4.9 The Home Secretary will have additional powers in relation to the codes of practice on preventing terrorist use of the internet and child sexual exploitation and abuse. The Home Secretary will be able to require the regulator to review the codes of practice on child sexual exploitation and abuse, and terrorist content and activity. This reflects the Home Secretary's responsibility for national security and the government's response to online child sexual exploitation and abuse.

4.10 To ensure proper parliamentary scrutiny of the codes, the objectives will be debated and voted on in Parliament under the affirmative resolution procedure. The individual codes, and any subsequent material amendments to them, will also be laid in Parliament and will be subject to the negative resolution procedure. Doing this will enable greater flexibility to respond to emerging threats and changing behaviours.

Promoting innovation

White Paper: *The White Paper proposed that the regulator should have a legal duty to pay due regard to innovation, to ensure competition within regulatory markets, and to help companies find more efficient ways of working with the regulator.*

Consultation responses and stakeholder engagement: *Engagement suggested that there is significant appetite for government influence and advocacy to support innovation.*

Final policy position: *Ofcom, as the independent online harms regulator, must already pay due regard to encouraging innovation and promoting competition in relevant markets when performing its duties, as set out in section 3(4)(d) of the Communications Act 2003. A comparable duty to pay due regard to promoting innovation in relation to online harms will be put in place by the new legislation*.

4.11 The White Paper set out that the regulator should have a legal duty to pay due regard to innovation, to ensure competition within regulatory markets and to help companies find more efficient ways of working with the regulator. Since the White Paper the government has undertaken additional work with industry and third sector stakeholders to understand how this duty can best be delivered.

4.12 The Communications Act 2003 establishes that Ofcom must pay due regard to encouraging innovation. A comparable duty to pay due regard to promoting innovation will also apply to Ofcom's implementation of the regulatory framework around online harms, and will be underpinned by a new statutory duty requiring Ofcom to publish information setting out how it will encourage innovation with regards to online harms.

Transparency

White Paper: *The White Paper set out that developing a culture of transparency, trust and accountability will be a critical element of the new regulatory framework. It stated that companies in scope will be required to publish annual transparency reports. These reports will include, for example, information about the prevalence of harmful content or activity on their services and what measures are being taken to address it.*

Consultation responses and stakeholder engagement: *Respondents to the consultation and stakeholders highlighted the importance of transparency in holding companies to account for enforcement of their own standards, and upholding freedom of expression. Industry respondents suggested that transparency requirements should be proportionate – noting that a 'one size fits all' approach was unlikely to be effective and could be costly to implement for smaller companies.*

Final policy position: *The future transparency reporting requirements are in line with the proposals set out in the White Paper. The future framework will be future-proof and proportionate, and will give Ofcom the flexibility to determine the specific information companies will need to provide.*

The Government Report on Transparency Reporting in relation to Online Harms: *The government established a multi-stakeholder Transparency Working Group, which includes representatives from civil society and industry. This group produced recommendations on the future transparency framework, which the government has set out in The Government Report on Transparency Reporting in relation to Online Harms published alongside the Full Government Response.*

4.13 The regulatory framework will improve transparency about the processes that companies have in place to keep users safe. This will help to ensure that Ofcom, users and civil society understand the decisions that companies are making and can hold them to account.

4.14 Transparency reporting will help empower users to make informed decisions about their online activity. By highlighting the steps that companies are taking to keep their users safe, these reports will help drive industry accountability and encourage action from companies.

4.15 Companies providing Category 1 services (see Part 2 for further details) will be required to publish reports containing information about the steps they are taking to tackle online harms on these services. The Secretary of State for Digital, Culture, Media and Sport will also have the power to extend the scope of companies who will be required to publish transparency reports, beyond Category 1 companies, by setting additional thresholds based on factors such as the functionalities and the audience of the service.

4.16 It is likely that the information that will be most useful to the regulator and to users will vary between different companies. To ensure that the transparency framework is proportionate and reflects the diversity of services in scope, the transparency reporting requirements will differ between different types of companies. Ofcom will consider companies' resources and capacity, service type and audience in determining what information they will need to include in their reports.

4.17 Furthermore, to ensure that the transparency reporting framework is agile and future-proof, the regulator will need flexibility in determining the specific information companies will need to provide. The legislation will set out a list of the types of information that the regulator may require companies to report on, relating to a number of areas.

4.18 An indicative list of the high level categories of information that companies might need to include in their transparency reports is set out in Box 17 below.

Box 17: What types of information will transparency reports cover?

- Information about the enforcement of the company's own relevant terms and conditions, which should reflect the regulator's codes of practice.
- Information about the processes that the company has in place for reporting harmful content and activity (including in relation to illegal harms), the number of reports received and the action taken as a result.
- Information about the processes and tools in place to address illegal and harmful content and activity, including, where appropriate, tools to identify, flag, block or remove illegal and harmful content and the processes that companies have in place for directing users to support and information.
- Information about the measures and safeguards in place to uphold and protect fundamental rights, ensuring decisions to remove content, block and/or delete accounts are well founded, especially when automated tools are used, and that users have an effective route of appeal.
- Where relevant, information about evidence of cooperation with UK law enforcement and other relevant government agencies, regulatory bodies and public agencies.
- Information about measures to support user education and awareness of online harms and strengthen users' media literacy, including through collaboration with civil society, small and medium-sized enterprises and other companies.
- Information about tools for users to help them manage harmful content and activity.
- Information about the process and steps an organisation has in place to assess risk of harm at the design, development and update stage of the online service.
- Information about other steps that companies are taking to tackle online harms and fulfil their obligations under the online harms framework, including to deliver a higher level of protection to children where a platform is likely to be accessed by children.

4.19 The indicative list has been informed by the recommendations in the first Government Report on Transparency Reporting in relation to Online Harms, published alongside this response. The Report sets out the recommendations produced by the multi-stakeholder Transparency Working Group on what transparency reporting should look like, both as part of the future regulatory framework but also in the interim period.

4.20 Where the regulator has determined that a company should report and set out what they will need to report on, the company will be required to do so or will face enforcement action. Companies will be required to publish their reports and make their reports accessible. The regulator will publish guidance to provide further clarity to companies on its approach.

4.21 The regulator will be responsible for producing an annual report of its own which will summarise key findings and insights from the reports that companies have produced and will highlight good practice. This report will play a vital role in helping users and parents understand the differences between online services and make informed decisions about which ones they use.

Information gathering and Investigation

White Paper: *The White Paper set out that the transparency, trust and accountability framework would be backed by information gathering powers, to enable Ofcom to assess companies' compliance with the duty of care and develop its understanding of the risk landscape.*

Consultation responses and stakeholder engagement: *Respondents to the consultation did not answer specifically on information gathering and investigation powers but highlighted the importance of transparency in holding companies to account for enforcement of their own standards. Many stakeholders recognised the importance of equipping the regulator with the powers needed to determine whether companies are fulfilling the duty of care and emphasised that these powers should be used proportionately.*

Final policy position: *The regulator will have powers to require additional information from companies to inform its regulatory activity, including additional powers to support investigations.*

Information gathering powers and powers to support investigation

4.22 The regulator's information gathering powers will play a crucial role in supporting its various regulatory functions. These powers will help the regulator build an in-depth understanding of the online harms landscape, prioritise its activity and oversee companies' compliance with the regulatory framework.

4.23 The regulator will have a broad power to require the information that it needs to carry out its functions. This will give Ofcom the flexibility to determine the specific information it requires.

4.24 This power will apply to companies in scope of the duty of care and, where necessary, to other organisations or persons who may have relevant information. The regulator will be required to take a proportionate approach in exercising its powers.

4.25 Ofcom will use information from a wide range of sources to help prioritise its investigation and enforcement activity. Alongside the information which companies have provided (in their transparency reports and in response to information requests) the regulator will also utilise user complaints data and publicly available information to help determine whether an investigation might be warranted.

4.26 The regulator will also have a number of additional powers to support its oversight and enforcement activity. Where there are reasonable grounds to suggest that a company may be non-compliant, Ofcom will have the power to enter companies' premises and access documentation, data and equipment in order to understand whether companies are taking sufficient measures to fulfil the duty of care.

4.27 Ofcom will also have a power to interview employees, which will allow it to develop further understanding of how the company is complying with the duty of care.

4.28 Finally, Ofcom will have the power to require a company to undertake, and pay for, a skilled person report on specific issues of concern. This power will be particularly useful on issues where external technical expertise is needed, for instance to validate the effectiveness of automated moderation systems. As with all its powers, Ofcom will be required to take a proportionate approach to the use of this power.

Researcher access to company data

4.29 To support research into online harms, and to help the regulator to prioritise its actions, Ofcom will be required to produce a report on the opportunities, challenges and practicalities of companies providing independent researchers with access to company data to support research into online harms.

4.30 As part of this Ofcom will produce best practice guidance for companies and researchers on how to approach it. In preparing this guidance, Ofcom will be required to consult a broad range of stakeholders, including companies, academics, the Information Commissioner's Office, the Centre for Data Ethics and Innovation, UK Research and Information.

Box 18: Ahead of the research activity that Ofcom will undertake, the Department for Digital, Culture, Media and Sport will deliver a comprehensive package to help inform and shape our work on online harms. This includes:

- an award of £2.6m between 2020-21 and 2021-22 from HM Treasury's Shared Outcomes Fund for a project to address current barriers to data sharing and to improve data interoperability to support innovation and competition in the detection of online harms;
- a phased study to investigate the feasibility of research to assess the drivers and impact of online harms;
- research into the impact on business and operational concerns surrounding the implementation of the UK-established video sharing platform regulatory regime;
- research that will consider the possible exclusion risks posed by age assurance solutions to vulnerable children;
- research that will consider the relationship between platform design and online harms.

User redress

White Paper: *The White Paper committed to ensuring measures are in place for users to seek redress, and consulted on a proposed super-complaints framework. It also noted that users would be able to alert the regulator to their concerns, and use regulatory decisions in legal proceedings.*

Consultation responses and stakeholder engagement: *Organisations overwhelmingly agreed that companies should have effective, accessible and transparent mechanisms for reporting harmful content and felt that current processes often fell short. They agreed that this process should start with reports directly to the service, and noted the importance of making these mechanisms accessible and prominent for all users, including children.*

Final policy position: *As detailed in Part 2, companies will be required to have reporting and redress mechanisms. The regulator will have oversight of these mechanisms. Ofcom will also establish a super-complaints function and user advocacy mechanisms to ensure it is understanding users' experiences, detecting issues early and addressing their concerns.*

4.31 As detailed in Part 2, all companies in scope will be required to have effective and accessible user reporting and redress mechanisms for the types of content and activity which they have to address as part of their duty of care. They will also be required to have mechanisms for users to report broader concerns about a company's compliance with its duties. The regulator will be able to access information about companies' reporting and redress mechanisms in the exercise of its statutory functions.

4.32 In addition to users being able to report their concerns to services, they will also be able to report their concerns to the regulator. However, the regulator will not investigate or arbitrate on individual cases. Allowing the regulator to do so would conflict with the principle of a systems and processes approach, and the number of potential complaints could overwhelm it. Instead, receiving user complaints will be an essential part of Ofcom's horizon-scanning, research, supervision and enforcement activity.

4.33 The government does not intend to establish an independent resolution mechanism, such as an ombudsman or certified Alternative Dispute Resolution scheme, for users to seek individual redress independently of companies. Such mechanisms are relatively untested in areas of non-financial harm. It is unclear how they would work in practice for online harms disputes, which centre on complex issues of safety and users' rights, or whether they would be valuable to users.

4.34 Establishing an independent mechanism for resolving disputes would not align with our overarching objective to ensure companies take more responsibility for their users' safety, and to improve users' trust in their processes. It could disincentivise cultural change within companies, and encourage companies to 'offload' difficult content decisions externally.

4.35 The government and the regulator will continue to assess evidence as the new framework comes into force. The Secretary of State for Digital, Culture, Media and Sport will undertake a review on the effectiveness of the regime 2-5 years after entry into force. This will offer an opportunity to re-assess whether the case for a statutory independent review mechanism is stronger, when the regulatory framework is better established.

Legal action by individuals

4.36 The regulatory framework will not establish new avenues for individuals to sue companies. However, the existing legal rights individuals have to bring actions against companies will not be affected. As outlined in the White Paper, the government expects legal action to become more accessible to users as the evidence base around online harms grows, and as regulatory precedent is established. Users will be able to use regulatory decisions that are publicly available as evidence in any relevant legal action they pursue.

Super-complaints

4.37 As proposed in the White Paper, a super-complaints function will ensure that there is an avenue for organisations representing users or those who are affected by harmful content and activity online (for example, victims of child sexual exploitation and abuse) to alert Ofcom to their concerns about systemic issues.

4.38 Under this function, Ofcom will accept super-complaints demonstrating substantial evidence of a systemic issue that is causing harm, or risks causing harm, to large numbers of users or specific groups of users. This will include those who may suffer disproportionately from online harms. Super-complaints will need to focus on the systems and processes that companies have in place, rather than any specific content issues. They will also need to focus on issues occurring across multiple in-scope services, as organisations can raise concerns about a single company's conduct through Ofcom's enforcement complaints processes. However, recognising the dominance of some services, super-complaints regarding one service will be admissible in exceptional circumstances.

User advocacy

4.39 Ofcom will also have a legal duty to establish ongoing mechanisms for user advocacy. These will ensure it understands the experiences of service users (including children) and others who are affected by harmful content and activity online (for example, victims of online child sexual exploitation and abuse), and that it can take action to address their concerns. It will also allow Ofcom to become aware of issues at an early stage before they can cause significant harm.

4.40 Ofcom will have discretion to determine appropriate user advocacy mechanisms, which may include expert panels, research, user panels or focus groups. This flexibility will ensure it is able to use the most appropriate methods for understanding users' concerns and experiences, and to encourage innovation in advocacy models. Ofcom will be required to report on its user advocacy work in its annual report to Parliament.

Enforcement

White Paper: *The White Paper set out that the regulator will have a range of enforcement powers to take action against companies that fail to fulfil their duty of care. It recognised that the powers must incentivise compliance and be used in a proportionate manner.*

Consultation responses and stakeholder engagement: *Stakeholder feedback expressed an overall preference for the regulator to begin its operations by supervising companies and supporting compliance through advice, and that any further enforcement measures should be used proportionately and following a clear process.*

Final policy position: *The principles and objectives underlying the enforcement proposals have not changed fundamentally, but the government has provided further details on what enforcement activity will look like. This includes refining the additional enforcement powers that the government consulted on. The most notable developments are in our approach to nominated representatives, senior management liability and business disruption measures.*

4.41 The approach to enforcement will aim to encourage compliance and drive positive cultural change. The regulator will support businesses to help them understand the expectations placed on them, and how the regulator's use of its enforcement powers will be proportionate. Ofcom will have a suite of enforcement powers available to use against companies who fail to fulfil the duty of care, or fail to put in place appropriate measures after being alerted to an issue. These powers will be comparable to those already used by Ofcom and other UK regulators. Ofcom will use its enforcement powers in line with its duties, including being proportionate, taking into account the level of harm and considering the impact on children.

4.42 The government recognises the need to balance effective enforcement with protecting the attractiveness of the UK as a tech sector, and also with users' rights. The regulator will strongly encourage compliance with the regime in the first instance and provide clear grounds for any intervention and escalation. The focus will be on ensuring that companies have compliant systems and processes in place, rather than on specific pieces of content.

4.43 The regulator's enforcement powers will include issuing directions for improvement and notices of non-compliance. Ofcom will have the power to issue sanctions in the form of civil fines up to £18 million or 10% of annual global turnover, whichever is higher. The fine limit is in line with the limits for those fines currently issued by Ofcom, the Financial Conduct Authority and the Competition and Markets Authority. As a last resort in cases of repeated or particularly egregious non-compliance, Ofcom will be able to take measures to disrupt a company's business activities in the UK, including blocking access in the most serious circumstances. If a company fails to fulfil their duty of care, the regulator may be able to pursue enforcement action against a parent company that wholly owns or controls the non-compliant company. Ofcom will take a proportionate approach to its enforcement activity and will be required to consult and publish guidance setting out how the powers will be used. Further details of the enforcement powers are set out in Box 19.

4.44 Alongside meeting their duty of care, companies in scope may also be required to provide transparency reporting, respond to information requests, use automated technology to remove illegal content (see paragraphs 2.54 to 2.70), notify the regulator in relation to the annual industry fee and pay the annual fee. Further details of these requirements are set out in Parts 2, 3 and 4**.** The regulator will be able to take enforcement action against companies that fail to comply with these requirements, using the powers set out in Box 19. In addition the regulator may also require information from third parties that are not in scope (see paragraph 4.24). The regulator will be able to issue fines against third parties that fail to comply. The regulator will be required to produce guidance setting out how it will use the enforcement powers in these circumstances.

Box 19: The regulator's enforcement powers:

- The power to issue directions and notices of non-compliance.

- Fines up to £18m or 10% of annual global turnover, whichever is higher:
    - The regulator will produce guidance on how penalties will be decided. The guidance will be based on the regulator's operating principles, including proportionality and the extent to which harm was caused to children.

- Business Disruption Measures, Level One:
    - The regulator will have the power to take measures that make it less commercially viable for a non-compliant company to provide services to UK users.
    - The regulator will have the power to require providers to withdraw access to key services. If providers do not comply, the regulator will be able to enforce through a court order.

- Business Disruption Measures, Level Two (serious failures of the duty of care):
    - The regulator will have the power to take measures that block a non-compliant company's services from being accessible in the UK, by requiring the withdrawal of services by key internet infrastructure providers (e.g. browsers, web-hosting companies, app stores, online security providers or Internet Service Providers).
    - This approach is technology neutral to encompass future changes to how the architecture of the internet functions.
    - The regulator will be required to obtain a court order for Level Two sanctions ahead of requesting a provider to block access to the non-compliant company's service in the UK, to safeguard freedom of expression online.

Enforcement in an international context

4.45 The White Paper set out that the regulatory regime will need to handle the global nature of online harms and be designed in a way to ensure the regulator can take action against companies without a legal presence in the UK. Ofcom should have powers to ensure a level playing field between companies that have a legal presence in the UK, and those who operate entirely from overseas.

4.46 It will be possible for the regulator to take enforcement action against any company, irrespective of where it is based in the world, if it provides services to UK users that are in scope of the online harms regime.

4.47 The enforcement powers have been designed to be able to be used against companies with and without a physical or legal presence in the UK. As other countries introduce similar legislation, international cooperation will become an increasingly important and effective tool for the regulator. The government expects the regulator to work with equivalent organisations internationally to help foster collaboration.

Nominated representatives

4.48 The White Paper proposed that companies should have nominated representatives in the UK or European Economic Area, to assist the regulator in taking enforcement action against companies based outside of these areas. While respondents acknowledged that this system would support the effectiveness of the proposed legislation, concerns were raised about the potential impact on business costs and operations. These would be particularly acute for smaller businesses. The government has decided not to proceed with this option. Ofcom may choose to request names of individuals through the notification process to act as a point of contact. Further details on the notification process are contained in the Funding Model section.

Senior Management Liability

4.49 The government also consulted on whether senior managers should be personally liable for failures to meet the duty of care. This emerged as an area of concern, with industry highlighting the risk of potential negative impacts on the attractiveness of the UK tech sector. Any sanctions for senior managers should support engagement with the regulatory framework. It will be crucial for the regulator to have access to reliable and timely information, so it can understand the impact of the regulation and how the duty of care is being met. Therefore the government will reserve the right to introduce criminal sanctions for senior managers who fail to respond fully, accurately, and in a timely manner, to information requests from the online harms regulator. This power, and the associated criminal penalties for failing to comply, will be consistent with Ofcom's existing information gathering powers. This power would not be introduced until at least two years after the regulatory framework comes into effect, based on a review of the impact of the framework. The sanction would be a last resort, only to be used if industry failed to meet their information sharing responsibilities. This approach balances industry concerns with many stakeholders' support for the proposal as a way to drive culture change.

Appeals

White Paper: *The White Paper set out that companies, and other individuals, will have the ability to seek judicial review of the regulator's actions and decisions through the High Court, to provide confidence that the regulator is acting fairly and within its powers.*

Consultation responses and stakeholder engagement: *The government consulted on whether there should be an additional statutory mechanism of appeal, who should be able to access this, and what the circumstances and standard for appeals through this route should be. Responses were broadly in support of a statutory mechanism in addition to judicial review, with the primary focus on it being affordable and accessible.*

Final policy position: *The government is now proceeding with the option of an additional statutory mechanism for appeal, as considered in the consultation. Appeals will be possible to an appropriate tribunal, on the basis of judicial review principles.*

4.50 The government consulted on whether there should be a statutory mechanism for appeal, in addition to the option of judicial review. Appeal mechanisms provide a route for the regulator's decisions to be challenged, which help make regulations more robust and fair. Enforcement decisions are a particular area where it will be important to have an accessible route of appeal, due to the novel nature of the regulations and the range of different companies in scope.

4.51 The government will ensure that, in addition to judicial review through the High Court, there is an additional statutory mechanism of review by designating an existing statutory body to review appeals. By using an existing statutory appeals body the regime will seek to save costs, and to reduce the financial burden on smaller businesses and interested third parties who wish to appeal decisions.

4.52 Appeals to an appropriate tribunal on the regulator's decisions will be on the basis of judicial review principles. This means that the tribunal will assess the legality of the decision and the process used to make it, rather than conducting a review of the merits of a decision.

4.53 The government recognises that consultation responses expressed a preference for a merits-based appeal process. The regulator will have the knowledge and expertise to determine the facts of the individual case. Rather than an appeals body seeking to gather and review these facts again, it will be best used to determine whether Ofcom has exercised its powers lawfully and fairly. The government is confident that an appeal using judicial review principles will offer both expediency, and appropriate levels of oversight of regulatory decisions, without undermining the regulator's decision making authority.

4.54 Any party with sufficient interest in the matter to which the application relates will be able to appeal Ofcom's decisions and sanctions, in line with judicial review standards. The government understands that a number of stakeholders will have an interest in the decisions of the regulator, including industry, civil society groups, and users. This approach is in line with international standards and is an important safeguard to ensure the regulator acts in accordance with its overarching principles and purpose.

Part 5: What part will technology, education and awareness play in the solution?

Summary

Consultation questions covered in Part 5:

- ❖ *What are the greatest opportunities and barriers for (i) innovation and (ii) adoption of safety technologies by UK organisations, and what role should the government play in addressing these?*
- ❖ *What, if any, are the most significant areas in which organisations need practical guidance to build products that are safe by design?*
- ❖ *Should the government be doing more to help people manage their own and their children's safety and if so, what?*
- ❖ *What, if any, role should the regulator have in relation to education and awareness* activity?

- The regulatory framework will be supported by an ambitious programme of practical support for the tech industry, which will put in place the guidance, tools and support needed to create safer online experiences for users. A renewed focus on media literacy will ensure users are better able to manage risks.

- The UK is a world-leading provider of ‘safety technology’ - products and services that help to deliver safer user experiences. The government will continue to support the growth of this sector, so that firms of all sizes have access to the technology they need to protect their users and comply with regulation.

- The government will develop a safety by design framework that will provide guidance for industry on how to build safer online products and services from the outset. It will encourage companies to actively consider the safety implications of their design decisions, and will be tailored to support companies with a range of digital skills and subject knowledge.

- The government and the regulator, working with industry, will both play a role in equipping users with the skills they need to manage risks online and critically appraise information. This will include publication by the government of a new online media literacy strategy.

- The government’s approach will include building upon the interactions between safety by design and media literacy, to promote the role of design in strengthening media literacy and improving user safety.

Safety tech market

White Paper: *The White Paper set out the government's ambition to position the UK as a world leader in safety technology. It proposed specific action to assess the online safety sector's capability and potential, and to explore how organisations can securely access training data to develop Artificial Intelligence solutions while ensuring that Artificial Intelligence use is safe and ethical.*

Consultation responses and stakeholder engagement: *The government has consulted with a wide range of stakeholders from across industry and civil society, to understand the potential for growth of the safety technology sector. Key themes emerging were the opportunities for government to:*

- *support a data infrastructure that enables greater innovation and competition in safety technology, for example by improving access to datasets that can be used for training Artificial Intelligence solutions;*
- *champion the emerging UK safety tech sector, including through growing international trade and improving sector access to funding sources;*
- *strengthen networks for collaboration within the safety tech sector and with the wider tech sector, and use insight from sector providers to inform policy development.*

Final policy position: *Since the White Paper, the government has published a detailed analysis of the UK safety tech innovation ecosystem, 'Safer technology, safer users: the UK as a world leader in Safety Tech', and announced a package of measures to work towards making the UK a world leader in safety technology. (See 'Upcoming measures' below).*

5.1 The government's aim is to ensure that all companies have access to the technologies and tools they need to support safer online communities. The White Paper set out that the government will work closely with the technology sector including industry, academia and civil society to make the UK a world leader in innovative technology solutions to prevent child sexual exploitation and abuse, terrorism and other harmful behaviours. The regulator will use its position to drive the development of new technologies and support the sharing of tools and best practice across companies.

5.2 Since the White Paper, the government has conducted a detailed study into the safety tech market. These findings, published in the 'Safer technology, safer users: the UK as a world leader in Safety Tech' report in May 2020,[44] demonstrate that UK safety tech providers are at the cutting edge of technology development, offering products that are helping to protect millions of users worldwide. This market is also increasingly interesting for investors; the sector has seen annual growth rates of 35% in recent years, with revenues predicted to exceed £1 billion by the mid 2020s. Internationally, UK companies have around 25% of global market share, and the sector employs around 1,700 Full Time Employees across the UK, including in regional hubs in London, Cambridge, Leeds and Edinburgh.[45]

[44] 'Safer technology, safer users: The UK as a world-leader in Safety Tech' UK Government, May 2020 (last viewed in November 2020)

[45] 'Safer technology, safer users: The UK as a world-leader in Safety Tech' UK Government, May 2020 (last viewed in November 2020)

5.3 The government has also supported the launch of the UK Online Safety Technology Industry Association (OSTIA), a collective voice for the safety tech sector, which will help to increase visibility of new innovations, new technology and best practice for online safety. In August 2020, the Department for Digital, Culture, Media and Sport and the Department for International Trade published a Directory of UK Safety Tech Providers,[46] designed to help open up export markets.

5.4 To drive further sector growth, the government has also worked with stakeholders to identify opportunities for government support, based on five key areas of focus:

- Promoting the industry in national and international markets;
- Strengthening mechanisms for collaboration within industry, and the public sector;
- Targeted investment - running growth programmes in areas of policy priority;
- Improving data infrastructure;
- Convening the industry to identify consistent standards and strategies.

Upcoming measures

5.5 The government will continue to explore a range of measures to support the rapid development of the safety tech market. These are set out in Box 20, below**.**

Box 20: New measures to support the growth of the safety tech sector

To support the further growth of the UK safety tech sector, the government will:

- Deliver the Safety Tech Innovation Network, the world's first forum for safety tech providers to collaborate and promote their work;

- Deliver a new £2.6m project to prototype how better use of data around online harms can lead to improved Artificial Intelligence systems, and deliver better outcomes for citizens;

- Organise a series of events, including a Safety Tech Unconference and Expo, to raise awareness and showcase the best of safety tech to potential buyers;

- Help to organise trade missions to priority safety tech export markets;

- Collaborate across sectors, including with the UK Online Safety Tech Industry Association (www.ostia.org.uk), to identify opportunities for innovation, adoption and promotion of safety tech;

- Explore ways in which best practices in online safety can be included in standards and guidance for buying, building and reusing government technology, such as the

[46] 'Directory of UK Safety Tech Providers' UK Government, August 2020 (last viewed in November 2020)

Technology Code of Practice;

- Develop a Safety Tech Sector Strategy, to guide future priorities for sector support.

Safety by design, media literacy and engaging with information

Safety by design

White Paper: *The government committed to developing a safety by design framework to make it easier for start-ups and small businesses to embed safety during the design, development or updates of online products and services.*

Consultation responses and stakeholder engagement: *Stakeholders expressed broad agreement and recognition that standards of safety are improved when organisations build in user safety at the design and development stage of their online products or services. It was felt that greater guidance was needed on this, particularly for smaller companies. Whilst it was agreed that user safety should be the priority of a whole organisation, currently there is a notable gap in resources targeted at product designers, product managers and developers. The responses highlighted the need for greater specificity on the objectives of a safety by design approach and for the safety by design framework to be capable of supporting companies with a range of digital skills and subject knowledge.*

Final policy position: *The government remains committed to supporting the safer design of online products and services. The government intends to develop a framework of guidance by Spring 2021 to provide support to product designers, product managers and developers to help them adopt a safety by design approach.*

5.6 The White Paper stated that the government would deliver a safety by design framework, providing guidance to companies on how to design safer online products and services for users, and especially children. Service and product design decisions can directly impact on the likelihood of harms occurring online. It is therefore crucial that companies consider design choices to prevent harm and improve the safety of users' online experience.

5.7 The government has set out that companies will be expected to provide a higher level of protection to their child users, under the duty of care. The White Paper outlined the importance of media literacy in empowering users to engage critically with what they encounter online. In response to this, the safety by design framework will encourage companies to strengthen users' media literacy through design, provide particular protection to children and also reflect the ongoing responsibility that services have to their users.

5.8 A safety by design approach can apply from the conception stage of a new business onwards. User safety must be considered when designing the functionality of an online product or service, but also applies to setting in place an organisation's objectives and culture to fully support a safety by design approach. Companies should consider the impact of their design choices at each stage of the design and development process. Examples of a safety by design approach include: default safety settings, clearly presented information, positive behavioural nudges and user reporting tools that are simple to use. Further practical examples of design interventions taken by companies during the COVID-19 pandemic in response to disinformation and misinformation content online are in Box 21, below.

5.9 The government's safety by design framework will contain:

- High level design principles to guide product design and development work;
- Practical guidance for implementing safer design choices and effective safety features;
- Examples of best practice and case studies on service design.

5.10 The consultation responses demonstrated that companies will have different requirements from the safety by design framework depending on their knowledge of user safety, online harms and service design. For example, some organisations are digitally capable but lack the knowledge to identify and mitigate user safety risks. Others are informed about online harms but have limited digital skills to implement effective changes.

5.11 The safety by design framework will be tailored to meet a range of different needs. The government will engage with companies of different sizes, capabilities, and sectors to develop and test it. The government will also work closely with technical experts, industry, academia and civil society to ensure the right approach is adopted.

5.12 All companies, whatever their size, need to have the right tools to pre-empt and mitigate the risk of misuse of their services. The safety by design framework will help designers and developers consider the safety implications of their design decisions and incorporate existing good practice into their products. This will support fulfillment of the duty of care, improve standards of user safety and strengthen users' media literacy. Responsibility for promoting safety by design will ultimately pass to the regulator.

Media literacy

White Paper: *In the White Paper the government committed to developing an online media literacy strategy for both adults and children. It also stated that industry and government have a shared responsibility to empower users to manage their online safety. It set out that the regulator will have oversight of industry activity and spend on education and awareness, and a responsibility to promote online media literacy.*

Consultation responses and stakeholder engagement: *While some respondents felt that the regulator should not have a role in education and awareness, others made a range of suggestions for how the regulator might take specific action. These included overseeing industry activity and spend; creating an evaluation framework for assessing education and awareness activity; and promoting awareness of online safety.*

Final policy position: *The government will publish its forthcoming online media literacy strategy in spring 2021.*

5.13 Internet users want to feel empowered to manage their own online safety. However, as the White Paper identified, many adults and children do not think there is adequate support in place to understand the risks and feel vulnerable online as a consequence.

5.14 The government recognises the vital role that education can play in supporting children and adults to navigate the online world safely. In England, the Department for Education has introduced the statutory relationships, sex and health education curriculum (from September 2020), alongside the computing curriculum (from September 2014). Both support children to navigate the online world safely.

5.15 The Department for Education has also brought in new national standards for essential digital skills that set out the skills needed to operate effectively in life and work; 'Being safe and responsible online' is one of the 5 skill areas. The department has introduced new digital skills qualifications up to Level 1 that are based on these new standards, alongside a new legal entitlement for adults with no or low digital skills to study these qualifications free of charge. The government welcomes action already taken by companies, in partnership with civil society, to develop education and awareness programmes for online users.

5.16 Despite this progress, more needs to be done to equip users with the skills they need to spot risks online, critically appraise information and take steps to keep themselves and others safe online. This includes supporting adults, including in their role as parents and carers.

5.17 To achieve this, the government made a number of commitments in the White Paper which recognise that industry and government have a shared responsibility to empower users to manage their online safety. The online harms regime will build on Ofcom's existing responsibilities and empower it to play an enhanced role in improving media literacy across the board.

Role of Ofcom in media literacy

5.18 Ofcom has an existing statutory duty to promote media literacy under section 11 of the Communications Act 2003, which is currently delivered through its media literacy research and online research programme 'Making Sense of Media'.

5.19 The online harms regime will build on this statutory duty and be designed so that Ofcom can:

- Promote greater understanding of the public's media literacy knowledge and skills through research, identifying key gaps and groups with the greatest need and ensuring the public has access to current information;
- Develop a greater understanding of how service design choices strengthen users' media literacy (see paragraphs 5.29-5.32 below);
- Develop and encourage others to develop educational initiatives which increase public awareness and online safety;
- Support and encourage the evaluation of media literacy initiatives, including service design choices and educational programmes, through the development and maintenance of a media literacy evaluation framework.

5.20 Ofcom will be able to undertake a range of initiatives when it identifies an area in which media literacy needs to be improved. This might include communications campaigns, piloting targeted interventions and providing training to key services in the community (e.g. support workers, community leaders). Ofcom will have independence to discharge its duties in this regard, although the government may have views on the regulator's priorities that the regulator should take into account in determining its work.

5.21 Ofcom will also play a role in overseeing industry activity and spending on education and awareness, as well as the impact of service design upon media literacy. This will be delivered through the transparency reporting framework (see Part 4). Certain companies in scope may be required to report on their education and awareness raising activity, to allow users to make informed choices about their online activity and understand the support offered by different services. It is also possible that companies may be asked to report on media literacy initiatives specifically for children.

5.22 Although Ofcom will oversee industry activity, it will not have the power to direct industry spend or activity.

Role of the government in media literacy

5.23 The government committed to publishing a new online media literacy strategy, after broad consultation with stakeholders. The strategy will ensure a coordinated approach to online media literacy education and awareness for children, young people and adults. The online media literacy strategy will be published in spring 2021.

5.24 The government has consulted widely with stakeholders on the proposed strategy. It has also undertaken a comprehensive landscape mapping exercise to identify what actions are already underway, and to shape the objectives of the online media literacy strategy. Alongside this, the government has considered research on the levels of media literacy among users, and evaluated the evidence base for media literacy interventions.

5.25 The online media literacy strategy will focus on supporting users in managing their privacy settings and their online footprint. It will help them think critically about the content they come across online, including disinformation and misinformation, and how the terms and conditions of services and moderating processes can be used to address harmful content. The strategy will also acknowledge the importance of action from industry in ensuring that service design strengthens users' media literacy skills.

5.26 The strategy is designed to deliver tailored outcomes for different groups. It is intended to promote greater understanding of media literacy for children and young people, balancing their enhanced digital skills with their increased vulnerability online. It will also support parents to improve their media literacy skills whilst caring for children, so they can better understand and prevent the risk of harmful activity online.

5.27 The strategy will complement Ofcom's work in media literacy, including Ofcom's Making Sense of Media programme, and existing initiatives. This includes, in England, the work the Department for Education is leading on ensuring that schools are equipped to teach online safety and digital literacy and the introduction of a new legal entitlement for adults with no or low digital skills to study essential digital skills qualifications free of charge.

5.28 The government is working closely with the Devolved Administrations and with the seven Mayoral Combined Authorities and the Greater London Authority, where adult skills funding is devolved to take account of their priorities and existing programmes of work in the online media literacy strategy.

Engaging with information - how safety by design and media literacy work together

5.29 Service design and a user's critical engagement with online content are closely connected. They can either work together positively to improve a user's safety and wellbeing or can interact in a way that disempowers a user. Media literacy is influenced, in part, by the design and interface of online services and products.

5.30 Online services and products can be designed in a way that limits the ability of users to engage critically with online content. For example, a user journey that allows the user to forward messages to an endless number of people risks limiting the user's ability to critically assess content, and leaves them more vulnerable to engaging with misinformation and disinformation online.

5.31 However, service design can be harnessed to support and encourage a user's critical thinking. Good behavioural nudges can be used to prompt a user when they are at risk of encountering or sharing content that is potentially harmful or incorrect. Fact-checked, trustworthy content can be clearly marked and users can be provided with tools to manage the content that they see.

5.32 The government's approach to safety by design and media literacy aims to promote and improve the impact that service design can have on strengthening users' media literacy skills. The safety by design framework (see 5.9-5.12 above) will provide organisations with practical guidance on how to design safer online services and products that empower users. As part of this role, Ofcom will develop a greater understanding of how service design strengthens users' media literacy skills. This dual approach will empower adult users to keep themselves safe online and ensure companies consider the impact of their design choices on user safety.

Box 21: Engaging with information: the role of design in strengthening media literacy

The COVID-19 pandemic brought the danger of disinformation and misinformation content online into sharp focus. In response to this, companies introduced new design interventions focused on strengthening users' media literacy. Nearly all the major social media services made technical changes to their products, including techniques to protect user safety online.

- YouTube continues to remove content which denies the existence of COVID-19 or contradicts the World Health Organization or NHS medical information.[47] The service also prohibits adverts published alongside content which promotes harmful health-related content, including anti-vaccination information.[48]

- Facebook has expanded its work with fact-checkers to continue addressing misinformation. In March 2020, Facebook displayed warnings on roughly 40 million posts related to COVID-19, based on 4,000 articles reviewed by independent fact checkers. When users saw the warning labels, 95% of the time these users did not go on to visit the original content.[49]

- Other services have also taken steps to improve users' ability to find relevant information and improve their safety during COVID-19. Twitter serves "Know the facts" messaging to users who search for virus related information on the service, which directs users to the National Health Service website. They also provide a prompt for users who search for 5G content, directing them to government and authoritative sources of information.

- Such measures demonstrate that small design changes can potentially have a significant impact on user behaviour online; in this case ensuring people can stay safe by being better informed over the content they view during the COVID-19 pandemic.

[47] 'COVID-19 Medical Misinformation Policy' YouTube, May 2020 (last viewed in November 2020)
[48] 'Advertiser-friendly content guidelines: Controversial issues and sensitive events' YouTube (last viewed in November 2020)
[49] 'Facebook post: Mark Zuckerberg' Mark Zuckerberg, April 2020 (last viewed in November 2020)

Part 6: How does the regulatory framework fit into the wider digital landscape?

Summary

- *The online harms regulation is a key component of the government's future work to harness the opportunities of digital technology.*
- *Beyond online harms, the government is developing proposals in a range of areas to improve online safety and security, support dynamic and competitive digital markets and promote our democratic values online.*
- *The government's digital strategy will set out how the government is bringing the different strands of activity together in one place, as well as putting digital tech at the heart of the response to COVID-19.*

Wider digital strategy

6.1 The government has announced that it will publish a new and ambitious digital strategy, which will recognise the increased importance of digital technology and data in our lives, and the crucial role it must play in the future. The strategy seeks to ensure that the UK maximises the benefits of a tech-led recovery to the COVID-19 pandemic.

6.2 Pro-innovation governance, including regulation that builds trust and certainty, will be a key component of a future strategy, supporting the government's wider response to harness digital opportunities arising in the digital age. The government is taking regulatory action in a range of areas - including cyber security, competition, and protecting quality journalistic content - to improve online safety and security, support dynamic and competitive digital markets, and promote our democratic values online. Within this, online harms regulation will be a key part of the government's ambition for the UK to be the safest place in the world to be online, while taking a proportionate approach that promotes innovation - for example by building up the safety tech sector.

6.3 The government will also ensure that its approach to governing digital technologies is streamlined and coherent. Many of the harms relating to digital technologies have common underlying drivers, such as market power and information asymmetry, and different government interventions can often target the same companies. For example, action to tackle online harms needs to be consistent with work to promote high quality journalism given their shared focus on online content, consumer education, and the role of social media companies. A holistic approach is therefore necessary.

Box 22: Examples of wider government regulatory action

A pro-competition regime for digital markets

- On 27 November 2020, the government announced it is establishing a new, pro-competition regime for digital markets. The new regime will include:
    - an enforceable code for digital platforms with substantial and enduring market power, which will promote competition in digital markets including those funded by online advertising and ensure the sustainability of high-quality journalism and news publishing
    - the establishment of a dedicated Digital Markets Unit (DMU) in 2021 to introduce, maintain and enforce a code of conduct.
- The announcement was included in the government response to the Competition and Markets Authority's recent market study into online platforms and digital advertising. The Competition and Markets Authority found that Google and Facebook have market power in search, social media and online advertising markets.
- The government accepted the six strategic recommendations in Unlocking Digital Competition (the Furman Review), and at the March Budget, commissioned a new Digital Markets Taskforce to advise on the design and implementation of the pro-competition regime. The Taskforce published their advice in December 2020 and the government will respond in due course.

Cairncross Review on sustainable journalism

- The Cairncross Review was commissioned by the Department for Digital, Culture, Media and Sport in March 2018 to examine the sustainability of high quality journalism in the UK. The Review put forward a set of recommendations to help secure the future sustainability of the press sector, focusing on issues surrounding tech platforms, digital advertising and public interest journalism.
- The government is supportive of the majority of recommendations, including the publication of a media literacy strategy, support for platforms to help users better identify the reliability and trustworthiness of news sources and the introduction of a 'news quality obligation' on platforms.
- The Review also identified an unbalanced relationship between online platforms and news publishers, with the potential to threaten the viability of news publishers' businesses. In response, the Review recommended the establishment of new codes of conduct aimed at rebalancing the relationship between publishers and platforms.
- The enforceable code aimed at promoting competition in digital markets announced on 27 November 2020 is consistent with, and delivers on the substance of the similar proposal put forward in the Cairncross Review into sustainable journalism.

International Context

The White Paper set out our ambition that the UK's approach to online harms can lead towards new, global approaches for online safety that support our democratic values, and promote a free, open and secure internet. The government recognises that the proliferation of harms online is an international problem and that international collaboration is important *to tackle it. The government is continuing to engage with other countries as the government develops our approach and shares our experiences.*

6.4 Countries around the world are grappling with how to make the internet a safer environment for users. The regulator will take an international approach, working with other international regulators, to ensure effective enforcement and promote best practice at a global level.

6.5 The government continues to engage with international partners to learn from their experiences and build consensus around shared approaches to tackling online harms that uphold our democratic values and promote a free, open and secure internet.

6.6 International collaboration remains vital. The government welcomes international, industry-led, multi-stakeholder initiatives – including initiatives supported by the UN and other multi-lateral bodies – such as the Global Internet Forum to Counter Terrorism, the WePROTECT Global Alliance, and wider initiatives such as the Freedom Online Coalition and the Technology Coalition Fighting Child Sexual Abuse.

6.7 The UK government is a member of the newly established Independent Advisory Committee to the Global Internet Forum to Counter Terrorism alongside the governments of Canada, France, Japan, New Zealand, Ghana, and USA, as well as representatives from civil society. In 2020 the Global Internet Forum to Counter Terrorism became an independent organisation with a stated mission to prevent terrorists and violent extremists from exploiting digital platforms.

6.8 The UK is also a signatory of the 'Christchurch Call to Action', which was launched in May 2019 to prevent the abuse of the internet as occurred in and after the Christchurch attacks in New Zealand.[50]

6.9 In March 2020 the UK, alongside its Five Country partners, launched the Voluntary Principles to Counter Online Child Sexual Exploitation and Abuse. These principles set out actions companies can take to combat online child sexual exploitation and abuse, and were developed in consultation with tech companies and non governmental organisations.[51]

6.10 Since the publication of the Online Harms White Paper, more countries have taken action to tackle harmful online content domestically.

6.11 Ireland's General Scheme for an Online Safety and Media Regulation Bill was published in January 2020. The proposed Bill will establish a framework for the regulation of online safety to tackle the spread of harmful online content. A new Online Safety Commissioner will form part of a Media Commission which will govern the framework through both binding online safety codes and compliance, enforcement and sanction powers.

[50] 'Christchurch Call: To Eliminate Terrorist and Violent Extremist Content Online' (last viewed in November 2020)

[51] 'Voluntary Principles to Counter Online Child Sexual Exploitation and Abuse' Five Country Ministerial (last viewed in November 2020)

6.12 Following a consultation, in May 2019 Australia's Office of the eSafety Commissioner published Safety by Design Principles as a guideline for businesses to assess, review and embed user safety into online services. The initiative intends to drive up the standards of online product development by providing a template for all businesses to improve the transparency of their systems and empower users to manage their own safety.

6.13 The European Commission has recently consulted on a Digital Services Act package, which will update liability and safety rules for digital platforms, services and products. The Act will propose new rules to increase the responsibilities of online platforms and information society service providers and reinforce the oversight over platforms' content policies in the European Union. The consultation on the proposed legislation, published in June 2020, sought stakeholders' views to better understand issues around harms in the digital environment.

6.14 The government will continue to work with our international partners to promote user safety online, strengthen a free, open and secure internet and build public trust in digital services.

Devolution

6.15 Internet law and regulation is a reserved policy area. The White Paper stated that "Internet services and their regulation is a reserved issue, therefore the government intends for our proposed framework to apply on a UK wide basis". However, the government is conscious that some of the harms that will likely be in the scope and some aspects of enforcement involve devolved competences.

6.16 The government is working closely with our colleagues in the respective devolved administrations, to ensure that such issues are taken into account.

6.17 In addition, the regulator will need to be able to operate in the devolved jurisdictions and ensure that devolved considerations are effectively built into their work. Ofcom already has a strong presence in all of the devolved administrations, and close working relationships with the devolved administrations. The government is working with both Ofcom's devolved offices and the devolved administrations to ensure a joined up approach that builds on their previous experiences.

Part 7: Conclusion and next steps

7.1 The development of the online harms regime represents an important step in the UK's strategy to create a coherent and pro-innovation framework for the governance of digital technologies. Proportionate and risk-based regulatory interventions, underpinned by strong institutions, will build user confidence in the digital economy and drive economic growth. The proposed regime also highlights that online safety is a shared responsibility between government, users and companies. It is critical to ensure that users can make informed decisions and have tools available to them to manage their online experience.

7.2 The consultation highlighted the urgent need for action to protect users, particularly children, from significant harm. Companies and user groups welcomed the government's intention to provide regulatory clarity and certainty. The responses to the consultation also emphasised the need to ensure that the scope of the regulatory framework is tightly defined and that it includes strong safeguards for users' rights online. The further details and changes to the policy position set out in this document reflect the feedback that the government has received since the publication of the Online Harms White Paper in April 2019.

7.3 The Online Safety Bill, which will give effect to the regulatory framework outlined in this document, will be ready in 2021.

Annex A

Guiding principles for the regulatory framework

The overarching purpose of the regulatory framework will be to improve user safety online, with a particular focus on illegal harms and the protection of children. The regulator and companies will be required to carry out their responsibilities under the framework in line with the following guiding principles:

- ➢ **Improving user safety**: taking a risk-based approach that considers harm to individuals.
- ➢ **Protecting children:** requiring higher levels of protection for services used by children.
- ➢ **Transparency and accountability**: increasing user awareness about incidence of and response to harms.
- ➢ **Pro innovation:** supporting innovation and reducing the burden on business.
- ➢ **Proportionality:** acting in proportion to the severity of harm and resources available.
- ➢ **Protection of users' rights online:** including freedom of expression and right to privacy.
- ➢ **Systems and processes:** taking a system and processes approach rather than focusing on individual pieces of content.

How the new regulatory framework will be delivered against the guiding principles

Improving user safety

- The duty of care will require in-scope companies to have appropriate systems and processes in place to improve the safety of their users.
- The regulator will issue codes of practice to outline the processes that companies need to adopt to demonstrate that they have fulfilled the duty of care.
- The government has issued interim codes of practice on terrorism and child sexual exploitation and abuse, alongside this response, due to the seriousness of these illegal harms.
- All companies in scope will have to tackle relevant illegal content and activity on their services.
- All companies likely to be accessed by children will have to prioritise the protection of children. These companies will need to put in place measures to keep children safe from harmful activity and prevent them from accessing age-inappropriate or harmful content.
- Companies providing Category 1 services will have to fulfil a duty of care towards adult users accessing legal but harmful content and activity on their services.
- Companies in scope will be required to have effective user reporting and redress mechanisms.
- The transparency framework will allow users to make informed decisions about which services they use.
- The regulator will be able to take enforcement action against any in-scope company providing services to UK users, irrespective of where it is based in the world.
- The regulator will promote education about online safety and the use of safety technologies, to empower users and tackle online harms. This will be particularly

important for users to be able to critically and independently manage their own risks around legal harms.

- The government's safety by design framework will set out clear principles and practical guidance for companies on how to build safer online products and services, thereby reducing the burden on the user to manage their own safety.

Protecting children

- The regulatory framework will legally require companies likely to be accessed by children to provide a higher level of protection for children, to take reasonable steps to protect them from accessing age-inappropriate or harmful content, and to protect them from other harmful activity. This includes being targeted by offenders (for example, in cases of child sexual exploitation and abuse).
- The differentiated approach includes a focus on keeping children safe online. All companies in scope will be required to assess whether their service is likely to be accessed by children, and if so to take steps to protect children on their services.
- The regulator will require companies' user redress mechanisms, where appropriate, to be suitable for and accessible to children.
- The government expects the regulator to prioritise the protection of children in its approach to enforcement action.
- There is an existing programme of work to help deliver the commitment to protect children online, ahead of the introduction of the regulatory framework. This includes the Information Commissioner's Office Age Appropriate Design Code, the interim code on practice on Child Sexual Exploitation and Abuse and our 'One Stop Shop' guidance.

Transparency and accountability

- All companies providing Category 1 services will be required to publish transparency reports, which will empower users to make informed decisions about which services they use.
- The Secretary of State for Digital, Culture, Media and Sport will have the power to extend the scope of companies who will be required to publish transparency reports, beyond those providing Category 1 services, if necessary.
- The regulator will oversee the implementation of clear and transparent user redress mechanisms. The regulator will make public the outcome of super-complaints to ensure transparency in its decision-making processes.
- A statutory appeals mechanism to challenge the regulator's decisions will ensure the accountability of the regulator and build trust and credibility in the regime.

Pro innovation

- Companies will receive support from the regulator to understand and comply with the regulatory framework in a proportionate and effective manner.
- The differentiated approach will mitigate the risk of disproportionate burdens on smaller businesses by narrowing the scope of companies that will have to comply with the duty of care, with regard to legal but harmful content and activity accessed by adults.
- Companies that are judged to be sufficiently low-risk will be exempt from transparency reporting requirements, reflecting the diversity of services in scope.
- Ofcom's funding model will introduce a high threshold for notification and the payment of fees. Many small and medium-sized enterprises will not need to notify or pay fees, reducing the burden.

- Potential sanctions for non-compliance will be proportionate to potential or actual harm caused and the size and revenue of the company.
- The government, supported by the regulator, will deliver a series of measures to support the rapid growth of the safety tech sector, to help ensure that companies have access to a range of tools to deliver safer user experiences. The regulator will also be required to pay due regard to promoting innovation.
- Exemptions will apply to online product and service reviews as well as ‘below the line’ comments. This will reduce the regulatory burden on many low-risk businesses who have a low degree of user interactions and user generated content.

Proportionality

- All companies will be required to take reasonable and proportionate action to improve the safety of users on their services, but the framework will minimise burdens, particularly on small businesses and civil society organisations.
- The regulator will take a deliberately risk-based, targeted and proportionate approach to ensure its activity reflects the size, severity of harm and capacity of companies.
- To ensure proportionality, the regulatory framework will establish differentiated expectations on companies for illegal and legal but harmful content and activity. Companies in scope whose services are likely to be accessed by children will be expected to have robust systems and processes in place to protect children.
- The regulator’s initial focus will be on those companies whose services give rise to the biggest and clearest risk of harm to users.
- Minimum thresholds will apply for transparency reporting and the reporting requirements will be proportionate to the type of service and risk factors involved.
- Ofcom’s funding model will introduce a high threshold for the payment of fees. The fees will be proportionate to the company’s global revenue and activity.
- The regulator will take enforcement action on an escalating scale, using its powers in a proportionate manner, and will not require nominated representatives.
- The regulator will have strong enforcement powers, including business disruption measures, to be used as a last resort where other interventions have failed to tackle the harm occurring on a service.

Protecting users’ rights online

- The framework seeks to protect users' rights online, particularly the rights to freedom of expression and privacy. By reducing the prevalence of abuse online, it seeks to enable more people to exercise their right to freedom of expression online, without fear of abuse or discrimination.
- To protect freedom of expression, the regulation will treat illegal and legal but harmful content for in-scope services differently.
- Legislation will include safeguards for media freedom, ensuring users continue to have access to reliable information.
- The regulation will not put new limits on online anonymity. The regulatory approach will by design address abuse online whilst protecting freedom of expression.
- Effective transparency reporting will help ensure moderation is well-founded, as the decisions services make on content removal and user appeals on content removal will have greater visibility.

- Escalating enforcement sanctions will avoid incentivising content takedown, with judicial oversight to safeguard the most severe sanctions like blocking. However, companies in scope will retain their existing legal liabilities for illegal content.
- User redress mechanisms will enable users to challenge content that unduly restricts their freedom of expression and to more effectively appeal content removal.
- Ofcom will accept super-complaints demonstrating substantial evidence of a systemic issue that is causing harm, or risks causing harm, including about limits on freedom of expression.
- The regulator's powers requiring the use of technology to proactively identify illegal content and activity will be subject to strict safeguards, ensuring any interference with users' rights to privacy is proportionate to the risk of harm. These powers will be used only where there are no alternative measures that are capable of achieving the same aim.
- The regulator will produce an annual report, including a statement on how users' rights are being protected, which will be laid before Parliament.
- Any party with sufficient interest will be able to appeal the regulator's decisions, which is an important safeguard for the protection of users' rights.
- The government's media literacy strategy will support users to think critically about information online, to manage their privacy and to report harmful content.

Systems and processes approach

- The new regulatory framework will focus on the wider systems and processes that services have in place to deal with online harms, taking a proportionate and risk-based approach.
- Ofcom will not investigate individual pieces of content or arbitrate on individual cases. It will instead consider whether in-scope companies put in place appropriate processes to identify and mitigate the risk of harm to their users.
- Codes of practices will set out the systems and processes that companies need to adopt to fulfil their duty of care. These could include:
    - Processes for accurately assessing the risk of harmful content and activity occurring on a company's services.
    - Appropriate governance systems for managing risk, including the involvement of senior personnel.
    - Content moderation approaches for different types of harmful content.
    - Tools to support users to manage harm.
    - Processes to allow users to report harmful content or activity and to appeal the takedown of their content.
    - Processes to understand the impact of online safety measures on freedom of expression and introduce appropriate mitigating measures.
- This list is not exhaustive. Companies will be expected to tailor systems and processes to the services they offer and regulatory expectations will be proportionate to the severity of the potential harm and resources available to companies.

HM Government

ISBN 978-1-5286-2312-4
CCS1220695430