

# Bird & Bird

## The UK's Proposed Online Safety Bill *Are YOU in scope?*

September 2021



Commissioned by

IMAGE ANALYZER

# Contents

|                                                       |    |
|-------------------------------------------------------|----|
| Executive summary                                     | 1  |
| Introduction                                          | 2  |
| Who has to comply with the proposed legislation?      | 3  |
| In-scope services                                     | 3  |
| Exemptions                                            | 3  |
| What are the obligations?                             | 5  |
| Duty of care                                          | 5  |
| In-scope content                                      | 5  |
| Tiered approach to obligations                        | 6  |
| What are the penalties / enforcement measures?        | 9  |
| Regulatory authority                                  | 9  |
| Ofcom's enforcement powers                            | 9  |
| What steps should companies be taking to comply?      | 11 |
| Measures relating to terms and conditions             | 11 |
| Measures relating to reporting and redress mechanisms | 11 |
| Measures relating to content moderation               | 11 |
| Measures to protect children                          | 12 |
| Governance measures                                   | 12 |
| Closing thoughts from Image Analyzer                  | 13 |

# Executive summary

On May 12th 2021, the UK government published its draft Online Safety Bill which will place a legal obligation on online platform operators and interactive website owners to swiftly remove illegal content.

The Bill also requires service providers to take down material that is not illegal but poses a risk of harm to site users and the wider public.

While it is tempting to think that impending legislation will only apply to the largest social networks, the new law will bring many more organizations into scope, even if they are based outside of the UK.

The government's 145-page draft Online Safety Bill includes proposals to categorize companies based on the number of people using their online platforms. Large social media sites are likely to be classed as 'Category 1' companies and will have greater obligations to moderate content uploaded by billions of users. 'Category 2' companies are likely to include any interactive community platform that promotes the sharing of user-generated content, such as travel sites, online gaming, dating, private messaging and retail sites.

Failure to adequately moderate content could result in companies being fined up to 10 percent of their global turnover or having access to their services blocked in the UK.

While we anticipate that there may be further changes to the draft legislation, we commissioned this whitepaper, written by legal experts from Bird & Bird LLP, to provide clarity to organizations that may find themselves in scope to comply with the UK Online Safety regulation, which is expected to pass into law by 2022.

In the whitepaper, Bird & Bird explains what the UK legal proposals cover, who will need to comply, how the UK government plans to categorize organizations, the penalties for non-compliance, and offers guidance on steps that organizations can consider taking in preparation for the new law being passed.

Although the prospect of additional online user safety legislation can appear daunting, there are many business benefits to be gained from creating a safer online environment for your service users, particularly when your website, game, or platform is designed to be used by children. If harmful content is allowed to appear and remain on your site, whether through oversight or overwhelm, there is a risk that your online community will become toxic and users will leave your platform.

Responsible content moderation, in compliance with the Online Safety Bill, will help your organization to maintain a positive online user experience, while protecting your brand integrity and revenues.

We hope you find the whitepaper helpful.

**Crispin Pikes**  
CEO

Image Analyzer  
[www.image-analyzer.com](http://www.image-analyzer.com)

# Introduction

- 1      Aimed at improving user safety online, building public trust in digital services, supporting innovation and driving digital growth, the UK government is set to introduce a new regulatory framework which addresses harmful content online.
- 2      The UK Department for Digital, Culture, Media & Sport published an Online Harms White Paper in April 2019, which launched a public consultation process to gather views on how the risks posed by online content should best be tackled. The government published its initial response to this White Paper in February 2020 and its final response on December 15th 2020.
- 3      The Draft Online Safety Bill (the “Bill”) was published on May 12th 2021 and will be now be subject to intense pre-legislative scrutiny before being put before the UK Parliament for legislative approval.
- 4      The regime will have a broad scope, both in terms of content to be regulated and companies caught by the rules, in these respects going even further than the EU’s proposals in the draft Digital Services Act. It is also a regime that will have real “teeth” – in cases of non-compliance, the proposed regulator, the UK Office of Communications (Ofcom) will have the power to impose fines of up to 10% of global turnover or withdrawal of access to services.
- 5      Companies are therefore well-advised to familiarise themselves with the draft legislation now – for those who feel strongly enough, opportunities still exist to contribute to the legislative process and, for those who don’t, the draft legislation offers some guidance as to early steps they can take towards eventual compliance.
- 6      This paper sets out the key aspects of the proposed online safety regime, set out in the Bill. First, it explains the scope of the Bill, including its broad territorial scope and how it will apply to providers of user-to-user services and search services, and which content on these services will be covered.
- 7      There follows a detailed summary of the duties imposed by the Bill, and how a multi-tiered approach will apply to duties of different service providers. This includes a breakdown of the duties relating to illegal content which will apply to all in-scope service providers, the duties relating to legal but harmful content affecting children, and the Category 1 duties relating to legal but harmful content affecting adults which apply only to certain service providers. This is followed by a summary of the key enforcement powers that the regulator, Ofcom, will have including their powers in relation to investigating and imposing fines and other penalties.

Finally, consideration is given to the steps which companies should consider taking now and once the regime comes into force in order to comply with their duties. Clearly the content of the Bill will likely go through some significant changes, before passing into law, and secondary legislation is required to provide further detail on key aspects, but the draft Bill contains sufficient insight to enable us to suggest some preparatory steps likely to be worth taking if you think your business will fall within scope.

# Who has to comply with the proposed legislation?

## In-scope services

- 8 The Companies will be within scope of the Online Safety Bill if they are within the territorial scope and provide either:
- user-to-user services, defined as internet services which host content generated, uploaded or shared by a user of the service, which may be encountered by another user or users of the service; or
  - “search services” which is defined as internet services which are or include a search engine.
- 9 The Bill will apply to companies which fall within the above definitions and their service has “links with the United Kingdom”. A service is defined as having links with the United Kingdom if:
- it has a significant number of users in the UK;
  - the UK is a target market for the service; or
  - the service is capable of being used by individuals in the UK and there is a material risk of significant harm to individuals arising from content present on the service or content which may be encountered in or via search results
- 10 One aspect of the Bill’s scope which has been particularly intensely debated to date is the seemingly long-armed territorial scope of the proposed law. In the above definition a company may fall within scope merely if its services are capable of being used by a UK user. However, the government appears to have paid heed to concerns voiced about the potential breadth of this test, evidenced by the addition of the need for there to be a “material risk of significant harm” posed by the service - which adds a qualification to the test, albeit a somewhat subjective one.
- 11 Examples of services likely to be in scope include, but are not limited to: 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.
- 12 Much has been said about the government’s decision to extend the regime to cover private online interaction, meaning private messaging services will be regulated. The definition of “content” in the Bill specifically includes private as well as public communications and there is no clear distinction made between obligations for companies providing such services as opposed to other services within scope. However, the Bill does contain some safeguards designed to protect user privacy and it is expected that the proposed regulator, Ofcom, will provide further guidance on what measures are likely to be appropriate in the context of private communications.

## Exemptions

- 13 Exemptions will apply where the risk of harm is sufficiently low that any regulatory requirements would be disproportionate. Exemptions will apply to specific services rather than organisations. The services which will be exempt are set out in Schedule 1 to the Bill and they include:

- Online business services used internally by organisations, such as intranets and enterprise cloud storage;
- Online services managed by educational institutions, where those institutions are already subject to sufficient safeguarding duties or expectations;
- Email communication, voice-only calls, and SMS/MMS;
- Services with limited functionality such that the only way in which users are able to communicate is by posting or sharing comments or reviews, or expressing a view on comments and reviews such as by way of a “like” or “dislike” button. Platforms likely to be out of scope as a result of this exemption include newspaper websites with comments sections and retail sites providing products or services, where the only relevant content on the site is customer reviews.
- Services playing a functional role in enabling online activity (such as internet service providers and web-hosting companies, VPNs, browsers, content delivery service providers, app stores and security software).

14 Although the services in the last category above fall outside of scope in terms of the duty of care to address harmful content, the government has made clear that, given their role in the internet ecosystem, they will be obliged to comply with orders to enable Ofcom to exercise its powers of enforcement against in-scope services (for example, facilitating website blocking orders).

15 Paid-for advertising largely falls outside of the scope of the regime. However, organic adverts and influencer marketing may fall within the regime as they are often indistinguishable from other forms of user generated content hosted by a service provider. It is expected that further guidance will be provided to help service providers identify their obligations in relation to advertising, a fast-evolving space where particular navigational assistance will be necessary.



# What are the obligations?

## Duty of care

- 16 Put at its simplest, the duty is to do what you can to prevent harm to users by putting systemic measures in place to a) identify and assess harmful content and b) remove or otherwise appropriately deal with it. The Bill repeats throughout its various provisions that obligations imposed on companies will be risk-based and proportionate. In determining liability, then, it appears that context - including size and reach of a service, its particular features, its intended/unintended audience and the nature of the potential harm - will be everything.
- 17 More specifically, the duty of care will oblige companies to take reasonable steps to prevent proliferation of illegal content, remove illegal content once identified, and ensure that children who use their services are not exposed to harmful content (even if not illegal). A further obligation, to take clear and consistent steps to tackle lawful but harmful content which may be seen by adults, will apply to a subset of in-scope services, as explained further below.

## In-scope content

- 18 The definitions of illegal and harmful content contained in the Bill raise some questions.
- 19.1 Illegal content is defined as content which constitutes an offence under UK law or in respect of which the provider reasonably believes that it amounts to such an offence. This has the result that providers may have a duty to remove content which in fact is not illegal at all – and places quite a burden on providers to make that judgment call. The government has already called out terrorist content and child sexual abuse and exploitation content as illegal, and in-scope services are expected to take particularly stringent measures to address these, and further “priority illegal content” will be identified by the government in secondary legislation. For smaller companies with limited resources, these categories of harm should be tackled as an absolute bare minimum. New criminal offences can be created at any time. One area, for example, being considered by the UK government currently is a new offence of anonymous online abuse. If this becomes a new offence, this would of course extend the types of illegal content to be addressed under the Online Safety Bill.
- 19.2 A further difficult judgment call presents itself in the form of the definition of content which is harmful to children/adults. Again, the government will identify in secondary legislation certain “priority” content (and, for children, “primary priority content”, expected to include pornography and violent content, and “priority content”). However, for non-designated content, it is left to the provider to determine whether it is harmful. The Bill has adopted a highly impressionistic definition for these types of content - essentially, where a provider reasonably believes that, due to the nature of the content, there is a material risk of direct or indirect significant adverse physical or psychological impact on a child/adult of ordinary sensibilities. A similar test exists for content, the nature of which may not itself pose a material risk of harm, but the fact of its dissemination may do – in this case the harm must however be direct (not indirect). The questions these tests raise for providers are obvious, but we have underlined those parts we consider likely to be particularly challenging for those tasked with deciding whether or not to take action against such content.
- 19 Certain harms have been identified as being excluded from the scope of the regulatory framework where there are existing regimes in place to regulate them. The excluded harms are those resulting from breaches of intellectual property rights, breaches of data protection legislation, breaches of consumer protection law and cyber security breaches or hacking. The government’s response in December 2020 also suggested that harms arising from fraud would fall out of scope, however in its press release when publishing the draft Bill, it announced that certain frauds, such as scams arising on online dating services, would now be covered.

## Tiered approach to obligations

21 The regime will introduce a tiered approach to the obligations on services in scope:

- General duties applying to all services in scope (separated into those for user-to-user services and those for search services) to address illegal content and assess the likelihood of children accessing their services;
- Additional duties applying to services likely to be accessed by children (again, separated for user-to-user and search) to address content harmful to children; and
- Further additional duties applying to “Category 1” services (a sub-set of user-to-user services deemed higher risk due to the number of users and/or functionality they offer) to address content harmful to adults.

### General duties applying to all services

22 All companies within the scope of the regime will need to fulfil their duties:

22.1 To carry out an illegal content risk assessment:

- This requires service providers to carry out risk assessments to identify, assess and understand the risk of users encountering illegal content and determine the level of risk of harm to individuals.
- Factors to be taken into account include user base (size, make-up, typical use), risk of content being encountered, risk of harm being caused as a result (and nature and severity of such harm) and how functionalities/the design of the service may increase or decrease such risk.
- Risk assessments will need to be carried within three months of Ofcom publishing a report and guidance on risk assessments, which it is required to do under the Bill. Assessments should be repeated if Ofcom significantly changes the risk profile for a particular type of service, or if the provider intends to make a significant change to the design and/or operation of the service.

22.2 To ensure safety from illegal content, including:

- A duty to take “proportionate steps to mitigate and effectively manage the risks of harm to individuals” resulting from illegal content.
- A duty to have in place proportionate systems and processes to minimise the presence and dissemination of “priority illegal content” and swiftly take down such content where the service provider is alerted to its presence. A provider could become aware of such content via a user notification/complaint, or due to its own actions when, for example, carrying out a risk assessment required under the legislation.
- A duty to include details of how individuals will be protected from illegal content within the service’s terms of service. Companies will be required to ensure that their terms of service in this regard are clear and accessible, including to children, and that they are applied consistently.

22.3 Relating to reporting and redress:

- This requires service providers to use systems and processes which allow users to easily report and/or complain about illegal content or a company’s non-compliance with its illegal content safety duties. Such systems and processes will need to be accessible, transparent and easy to use.



- Companies will also need to have in place procedures by which users can make complaints relating to actions taken by the service provider in relation to content they have generated or shared, for example where a user believes that their content has been wrongfully removed or their use of the service has been wrongfully restricted.
- Ofcom will be charged with producing Codes of Practice outlining requirements for reporting and redress, which are expected to cover issues such as transparency, accessibility, communications with users, signposting and appeals.

#### 22.4 Relating to record-keeping and review:

- Providers will be required to keep appropriate written records of any risk assessment carried out in accordance with their duties under the regime, as well as to keep records of any steps taken in order to comply with their duties.

#### 22.5 To have regard to freedom of expression/user privacy:

- The obligation in this respect is merely to “have regard” to freedom of expression/right to privacy when designing systems and processes to tackle illegal content. Concerns have been expressed that, in a balancing act between safety “duties” and a requirement to “have regard to” other rights, providers will be incentivised to pay only lip service to such rights and err on the side of caution by removing content to comply with safety duties.

#### 22.6 To assess the likelihood of children accessing their services:

- Every provider must carry out this assessment and keep a record of it. It will only be considered “unlikely” that children could access the service if the provider can show that the systems/processes they have in place mean that it is “not normally possible” for children to access the service or any part of it.
- This assessment must be carried out within 3 months of Ofcom publishing guidance on the same, and whenever the provider makes a significant change to the design/operation of the service, or in the face of evidence of decreased effectiveness of its systems to prevent child access or of increased access by children to the service.

### Additional duties for services likely to be accessed by children

#### 23 If the conclusion of a service’s assessment of this likelihood is that children may access their services, then they will also need to fulfil duties:

##### 23.1 To carry out a children’s safety risk assessment:

- As with illegal content, this is to identify and assess the risk of children of different age groups encountering content which is considered to be harmful to children (assessment must be made in relation to each of (a) primary priority content (b) priority content and (c) non-designated content)

##### 23.2 To ensure children’s safety from content harmful to children, including duties to:

- take proportionate steps to mitigate and manage the risks of harm to children of different age groups as identified in the children’s safety risk assessment;
- take proportionate steps to mitigate the impact of harm to children from such content;
- in relation to “primary priority content” which is harmful to children, use proportionate systems and processes to prevent children from encountering such content; and
- in relation to other content that is harmful to children (whether priority content or non-designated harmful content), protect children in at-risk age groups from encountering such content.

### 23.3 Relating to child-friendly reporting and redress:

- These obligations are similar to those for illegal content, but of course all reporting and redress mechanisms are expected to be even more accessible/user-friendly, bearing in mind children of different age groups need to be able to take advantage of them.

### Additional duties for Category 1 services

24 Category 1 providers will have additional obligations under the new regime. The precise thresholds for who falls into this Category have not yet been announced. Two factors will be relevant to the Category 1 designation, with the threshold conditions for each to be set out in secondary legislation:

- The size of a service's audience - harm is more likely to occur on services with larger user bases, for example due to rapid spread of content and 'pile-on' abuse.
- Functionality - certain functionalities, such as the ability to share content widely or contact users anonymously, are more likely to give rise to harm.

25 Ofcom will designate in-scope services as Category 1 or 2 in a register of service providers, according to the thresholds set by the government. If a company falls within Category 1 it will need to fulfil its duties:

25.1 To carry out an adults' safety risk assessment;

- Again, to identify and assess risk of adults encountering content which is harmful to adults on the service (assessment to be made against (a) priority content of this type and (b) non-designated content of this type).
- Where non-designated harmful content is identified in the risk assessment, the provider has a duty to notify Ofcom of this.

25.2 To ensure adults' safety from content harmful to adults, including duties to:

- specify in the terms of service the approach to priority content that is harmful to adults and content that is identified as harmful in the most recent risk assessment; and
- ensure that those terms of service are clear, accessible and applied consistently.

25.3 To protect freedom of speech (and privacy), including specific duties to protect journalistic content and "content of democratic importance":

- These specific duties (which go beyond those placed on Category 2 providers to "have regard" to freedom of speech and privacy when taking action against illegal content) have been included to address concerns raised about the potential "chilling effect" over-removal would have on freedom of speech online;
- When deciding on safety policies and procedures, Category 1 providers have to carry out, and publish, an assessment of the impact upon freedom of expression and user privacy. It must then specify in its Terms of Service the steps it has taken to address this impact.
- Providers must also "take into account" the importance of freedom of expression of (a) content of democratic importance (content intended to contribute to democratic political debate in the UK); and (b) journalistic content (defined relatively broadly) in determining what action they take in relation to such content.

25.4 To have reporting and redress measures in place – as for illegal content and content harmful to children.

# What are the penalties / enforcement measures?

## Regulatory authority

- 26 Under the proposals, Ofcom – the existing communications and broadcasting regulator in the UK – will be appointed as the independent regulator in relation to the new regime. Ofcom will have a central role in the new framework, including:
- Deciding which services will fall into Category 1 and maintaining a register of the categorised services.
  - Conducting and publishing risk assessments to identify and assess the risks of harm presented by services under the regime. This will involve developing risk profiles for different kinds of regulated services.
  - Consulting on and drafting Codes of Practice for providers on how to fulfil their duty of care. At least one Code of Practice will be issued in relation to safety duties relating to each of (a) Child sexual exploitation and abuse (CSEA) content; (b) terrorist content; (c) other illegal content; (d) content harmful to children; and (e) content harmful to adults. Other Codes named in the Bill will relate to reporting and redress, content of democratic importance, and journalistic content. The significance of these Codes should not be underestimated – compliance will be assessed against their content. Service providers may take alternative steps to address risk to those set out in the Codes of Practice, but if they do so, they will have to be able to demonstrate to Ofcom that those steps achieve the Online Safety Objectives (set out in clause 30 of the Bill) and safeguard freedom of expression and user privacy.
  - Transparency reports - Ofcom will be responsible for determining which service providers should publish annual transparency reports, what information should be contained in those reports (from a long menu of options set out in the Bill) and for reviewing and assessing their content.
  - Conducting investigations into any suspected non-compliance. This can be initiated by Ofcom itself, or in response to a “super-complaint” (where a “qualified entity” can raise a complaint about systemic failings by one or several service providers on behalf of users). Ofcom has been given extensive powers under the Bill to require information, conduct interviews, and enter premises to inspect services if necessary.
  - Making enforcement decisions. The Bill gives Ofcom a range of tools to encourage compliance/penalise non-compliance. We describe these in more detail below.

## Ofcom’s enforcement powers

- 27 The penalties that Ofcom can hand down to enforce compliance under the regime are wide-ranging and heavy-hitting. Ofcom’s enforcement powers include the following:
- “Technology notices” can be issued to companies which Ofcom considers as not doing enough to comply with their safety duties in relation to illegal and harmful content, and particularly in relation to terrorism and child exploitation and abuse. Such notices are aimed at requiring the company to do more to tackle illegal content and can require the company to implement the use of specific technologies to combat such content. Ofcom will be required to publish guidance about how it proposes to exercise its powers in relation to technology notices.

- Enforcement notices will be issued where Ofcom considers that a company has failed or is failing to comply with the requirements of the regime. There will be a two-tiered approach, with Ofcom first issuing a provisional notice which, if the company fails to comply with its requirements in a specified period of time, will be followed by a confirmation decision.
- Such decisions can include financial penalties and/or mandatory requirements enforceable by an application to the court for an injunction. Ofcom will have the power to issue hefty fines of up to £18 million or 10% of the offending company's annual global turnover, whichever is greater.
- It will also have the power to order withdrawal of service, split into two levels of measures. Level 1 measures are "service restriction orders" which require providers of ancillary services, such as search engines, promoters, advertisers and payment processing companies, to take steps such as withdrawing any promotion or links to the service to which the order relates. Level 2 are "access restriction orders" which would require ISPs/app stores to cut off access to the service. This measure can only be used where other measures have failed or are unlikely to be sufficient to prevent significant harm.
- The Bill also reserves a power to introduce a senior management criminal liability regime, which the government may do, without the need to pass further primary legislation, if it feels that service providers are not taking their duties seriously enough.

# What steps should companies be taking to comply?

28 The Online Safety Bill has a long way to go to get through pre-legislative and legislative scrutiny before it comes into force. The Ofcom Codes of Practice will contain the real detail on how service providers should fulfil their duties, and these too of course are not yet drafted. Despite this, we identify below steps in-scope companies may wish to consider now to get ahead of the curve on eventual compliance:

## Measures relating to terms and conditions

29 It appears that the key guiding principles when it comes to terms of use are that such terms should be clear, accessible, easy to use and applied consistently. Consider a review of your terms of service with a view to:

- Ensuring terms are not overly long or complex (particularly where likely to be read by younger or vulnerable users);
- Ensuring terms are located sensibly on a site/app and in a format in which they will best present (bearing in mind, for example, mobile device restrictions);
- Encouraging engagement with terms and conditions, such as periodic prompts to users to read relevant parts; and
- Clarifying what content is unacceptable on the service and what steps will be taken to address this.

## Measures relating to reporting and redress mechanisms

30 Ensuring that the process by which users can notify service providers of valid concerns relating to illegal content and lawful but harmful content is as easy as possible seems likely to be a core component of the expected online safety programme. Consider:

- Making reporting and flagging mechanisms transparent, user friendly and easy to find;
- Clarifying in your policies the likely actions you may take in response to a report/complaint and on what timescales; and
- Particular tweaks you could make if children are likely to use your service, to ensure these mechanisms can be understood by individuals of any age.

## Measures relating to content moderation

31 The government has indicated that service providers will be encouraged to take different moderation approaches for different types of content, with the processes which should be adopted to be set out in Ofcom's codes of practice. An increase in automated content moderation seems an inevitable part of the compliance solution. Consider:

- Reviewing the current balance between human and automated content moderation for your service and how that might be altered to best accommodate a significant rise in moderation duties.
- Training processes in place for human moderators and whether they adequately address the various factors that the Bill indicates need to be taken into account (e.g. content of democratic importance/journalistic content if applicable).
- Any steps you could put in place to record, monitor and evaluate the effectiveness of moderation systems. This should include consideration of the amount of legitimate material that was incorrectly removed.

## Measures to protect children

32 The UK government has made clear that Ofcom will be charged with taking a strict approach to non-compliance in relation to children's online safety, and it is therefore crucial for in-scope services to get this aspect of the duty of care right:

- As a first step, carry out an assessment as to whether, as a result of systems/processes you have in place, it is not normally possible for children to access your service. If it is possible, then you will need to consider protection measures.
- Whilst age verification is probably not a blanket solution, appropriate use for certain parts of services, or certain functionalities, for instance, is likely to serve you well.
- Consider other means of age assurance – for example, use of algorithms to analyse patterns of behaviour/use of language by users on services – and how and where those might be appropriately deployed on your services.
- Consider rating or tagging systems for material which is not suitable for children or appropriate use of parental control systems.
- Consider also steps which might reduce the impact of harmful content on users (children and adults alike) and encourage users to keep themselves safe online. Examples of steps service providers may take include: labelling or deprioritising content; signposting towards support; or prompting users to redirect to alternative sources when they are at risk of encountering or sharing potentially harmful or misleading content.

## Governance measures

33 Service providers are likely to be encouraged under the new regime to create a culture of user safety and adopt appropriate governance systems for managing risk. This may include:

- Ensuring that senior management are involved in risk management governance;
- Considering appropriate training for product engineers on the government's Safety by Design framework and other guidance; and
- Engaging with third party content moderation providers, charities, NGOs and other specialists to ensure that senior managers and employees involved in content design and moderation are better educated on the risks of illegal and harmful content.



# Closing thoughts

We hope that the UK Online Safety Bill whitepaper provided you with a better understanding of the impending legislation and whether your organization is in scope to comply.

Where user comments and uploads to discussion boards, internet forums, gaming sites and blogs used to be managed by a webmaster, or team of moderators, the sheer scale of user-generated content enabled by social media, smartphones and increased connectivity, have made it increasingly difficult for human moderators to stay on top of the task.

Let's just consider some social media statistics for a moment:

|               |                                                                                                                         |
|---------------|-------------------------------------------------------------------------------------------------------------------------|
| <u>300M</u>   | Of the 300 million posts to Facebook each day, 35.73% are photographs and 15.09% are videos.                            |
| <u>95M</u>    | 95 million photographs are uploaded to Instagram daily.                                                                 |
| <u>350K</u>   | 350,000 hours of video are streamed on Periscope every 24 hours.                                                        |
| <u>500HRS</u> | 500 hours of video are uploaded to YouTube every minute.                                                                |
| <u>456K</u>   | There are 456,000 Tweets a minute.                                                                                      |
| <u>950YRS</u> | It has been estimated that it would take someone 950 years to watch all of the Snaps shared on SnapChat every 24 hours. |

A major challenge facing organizations today is that a small percentage of these posts are at best offensive, at worst abhorrent and likely to drive people off your site for good. As you've just read, the Online Safety Bill will legally oblige online service providers with UK users to swiftly remove illegal and harmful content.

By 2022 InterDigital/Futurescope anticipate that videos will make up 82% of consumer internet traffic. When people abuse your online service to upload images or live-stream video depicting or inciting graphic violence, illegal pornography, or child abuse, this material pollutes online communities and spills over into real world harms. This is what the UK Online Safety Bill sets out to tackle.

To keep illegal and harmful content off their sites, the largest social media companies have outsourced their

content moderation to contractors who employ many thousands of content moderators who review and flag the most toxic content for removal. However, even these huge, dedicated teams struggle to keep up with the sheer volume of content being uploaded.

For smaller websites and fledgling online communities, managing content moderation as they attract more users and scale their operations has become a strategic issue for the C-suite. Millions of daily posts can no longer be checked by humans alone. Backlogs can result in severe delays in removing harmful content, risking user safety and penalties for non-compliance with the new online safety legislation.

If your organization is struggling to protect your online communities from harmful visual material so that you can create a positive user experience and comply with impending laws, we are here to help you with our award-winning AI-powered visual intelligence technology, which can stop illegal and harmful images and videos from being uploaded to your site, leaving only the more nuanced visual content for human review.

To discuss your automated content moderation needs and discover why leading brands are using our technology, please visit [www.image-analyzer.com](http://www.image-analyzer.com) to request a demo, or contact me.

Sincerely,

**Crispin Pikes**  
CEO

Image Analyzer  
[crispin.pikes@image-analyzer.com](mailto:crispin.pikes@image-analyzer.com)

## IMAGE ANALYZER

Image Analyzer provides artificial intelligence-based content moderation technology for image, video and streaming media, including live-streamed footage uploaded by users.

Its technology helps organizations minimize their corporate legal risk exposure caused by employees or users abusing their digital platform access to share illegal or harmful visual material. Image Analyzer's technology has been designed to identify visual risks in milliseconds, including illegal content, and images and videos that are deemed harmful to users, especially children and vulnerable adults.

Founded in 2005, Image Analyzer holds various patents across multiple countries under the Patent Co-operation Treaty.

The company is a member of the Online Safety Tech Industry Association (OSTIA).

---

## Bird & Bird

[twobirds.com](http://twobirds.com)

Abu Dhabi & Amsterdam & Beijing & Bratislava & Brussels & Budapest & Copenhagen & Dubai & Dusseldorf & Frankfurt & The Hague & Hamburg & Helsinki & Hong Kong & London & Luxembourg & Lyon & Madrid & Milan & Munich & Paris & Prague & Rome & San Francisco & Shanghai & Singapore & Stockholm & Sydney & Warsaw & Satellite Office: Casablanca

The information given in this document concerning technical legal or professional subject matter is for guidance only and does not constitute legal or professional advice. Always consult a suitably qualified lawyer on any specific legal problem or matter. Bird & Bird assumes no responsibility for such information contained in this document and disclaims all liability in respect of such information.

Bird & Bird is an international legal practice comprising Bird & Bird LLP and its affiliated and associated businesses.

Bird & Bird LLP is a limited liability partnership, registered in England and Wales with registered number OC340318 and is authorised and regulated by the Solicitors Regulation Authority. Its registered office and principal place of business is at 12 New Fetter Lane, London EC4A 1JP. A list of members of Bird & Bird LLP and of any non-members who are designated as partners, and of their respective professional qualifications, is open to inspection at that address.

49377090.1

# Bird & Bird

## EU Proposal for a Digital Services Act *Are YOU in scope?*

September 2021



Commissioned by

IMAGE ANALYZER

# Contents

|                                                                               |    |
|-------------------------------------------------------------------------------|----|
| Executive summary                                                             | 1  |
| Introduction                                                                  | 2  |
| Who has to comply?                                                            | 3  |
| Territorial scope                                                             | 3  |
| A multi-layered approach                                                      | 3  |
| What are the obligations?                                                     | 5  |
| Over-arching provisions                                                       | 5  |
| Obligations that apply to all service providers                               | 6  |
| Obligations that apply to Hosting Services,<br>Online Platforms and VLOPs     | 6  |
| Obligations applying to Online Platforms and VLOPs                            | 7  |
| Obligations that apply to VLOPs only                                          | 9  |
| What are the penalties / enforcement measures?                                | 10 |
| Competent authorities                                                         | 10 |
| Sanctions                                                                     | 10 |
| What steps should companies be taking to comply?                              | 12 |
| Status under the EU Digital Services Act                                      | 12 |
| Use of automated content moderation                                           | 12 |
| Participate in the development of voluntary<br>industry standards             | 12 |
| Advertising contracts and practices                                           | 13 |
| Updating terms and conditions                                                 | 13 |
| Proactive monitoring of decision-making processes<br>and terms and conditions | 13 |
| User reporting functionality and complaints procedures                        | 13 |
| Closing Thoughts                                                              | 14 |

# Executive summary

In December 2020, the European Commission published its proposals for the European Union Digital Services Act (EU DSA), which aims to create a safer online environment for EU citizens by legally requiring online service providers to swiftly remove illegal content from their platforms.

As a provider of content moderation technology to multiple global organizations, and to assist organizations in making the necessary preparations to comply with the EU DSA, Image Analyzer commissioned the lawyers at Bird & Bird to write this whitepaper to provide early insights into the legal obligations contained in the EU DSA proposals.

The lawyers at Bird & Bird have scrutinised the proposals and presented the key points of the EU DSA, to allow organizations to understand their future obligations and potential legal exposure under the new legal regime. An important point for readers to note is the territorial scope of the proposed legislation. Any online service provider with a significant number of users in the EU will potentially fall within scope. As Bird & Bird writes in the whitepaper, *“The DSA has a broad extra-territorial scope irrespective of intermediaries’ places of establishment or residence. Based on a ‘substantial connection’ test, the law will apply to intermediaries based outside of Europe that provide their services within the EU’s internal market and have service users whose place of establishment or residence is within the EU.”*

To avoid placing a disproportionate legal burden on smaller organizations and stifling start-up innovation, the EU DSA proposes a 4-tiered approach with organizations meeting incremental legal obligations according to the type of service they provide.

Regardless of their tier, all organizations within scope will have to comply with Take Down Orders relating to illegal content on their platforms.

While we anticipate further updates will be required as the proposals are finalised, this whitepaper is designed to provide clarity on whether your organization will be in scope to comply with EU DSA and, if so, which of the four categories will apply to your organization and what your level of legal obligation is likely to be. We hope you find the contents helpful and that by providing this whitepaper now, your organization can be better prepared to comply with the EU DSA when it becomes law.

If you need assistance with preventing illegal content from being uploaded to your online platform, please get in touch.

**Crispin Pikes**  
CEO

Image Analyzer  
[www.image-analyzer.com](http://www.image-analyzer.com)



# Introduction

- 1 The aim of the EU's Digital Services Act (the "DSA") is to ensure a safer online environment for EU citizens by creating a modern and harmonised framework for the handling of illegal online content, products, services and activities on intermediary services. In doing so, the proposal incorporates longstanding protections for online intermediaries from liability in relation to content published by third parties and seeks to protect the fundamental rights of users of online services.
- 2 The DSA was published by the European Commission in December 2020 in the form of a Proposal for a Regulation. Discussions on the draft Proposal are ongoing in the European Parliament and among EU Member States in the Council. Both institutions are expected to put forward amendments to the draft text. It will take a minimum of 18 months before the EU institutions reach a final consensus on the proposed DSA and it is formally adopted and implemented by EU Member States. Depending on the eventual speed of the negotiations, the DSA is not likely to come into force before 2023.
- 3 The DSA in its current form contains a number of headline-worthy provisions, including requirements placed on intermediaries to strengthen action against illegal content on their platforms, a broad extra-territorial scope and the power for the European Commission to impose significant penalties for non-compliance. In the most serious cases, these penalties include fines of up to 6% of the global turnover in the preceding financial year for the largest online platforms and as a last resort the Commission can ask a court for a temporary suspension of a particular service.
- 4 This paper explains the core aspects of the proposed DSA which intermediary service providers offering their services to users in the EU need to be aware of. First, it assesses the extra-territorial scope of the proposed DSA, including how it will apply to service providers who do not have an establishment in the EU, and the proposed DSA's multi-layered approach to platform regulation, which imposes increasing obligations on service providers depending on the breadth of services they offer.
- 5 A detailed breakdown of the obligations set out in the proposed DSA follows, by reference to the different categories of service provider identified within the draft legislation, before a summary of the proposed regulatory/enforcement framework, including who will be the competent regulatory authorities and the level of fines they will be permitted to impose.
- 6 Lastly, consideration is given to the practical steps service providers can be thinking about and taking now, in order to ensure they are in the best position to demonstrate compliance with the DSA when it eventually comes into force across the EU. Clearly, we are a long way off a final version of the legislation, but the draft DSA provides an indication of the direction of travel which enables us to point to certain areas of preparation worth focussing on if you think your organization will fall within scope.

# Who has to comply with the proposed legislation?

## Territorial scope

- 7 The proposed DSA has an extra-territorial scope. It will apply to all intermediary service providers irrespective of their place of establishment or residence, if they provide their services within the EU's internal market (provided the recipients of their service have their place of establishment or residence within the EU).
- 8 Whether services are provided within the EU will be determined by a “substantial connection” test. A substantial connection will be deemed to exist where the service provider either:
- has an establishment within the EU; or
  - has a significant number of users in one or more Member States; or
  - targets activities towards one or more Member States.
- 9 The targeting of activities towards Member States will be determined on the basis of all relevant circumstances, including factors such as the language or currency used in that Member State; the possibility of ordering products to, or receiving services in, that Member State; or using a top level domain of that Member State. Other factors could be the availability of an application in the relevant national application store; the provision of local advertising or advertising in the language used in that Member State; or the provision of customer services in the language used in that Member State.
- 10 The DSA also provides that a substantial connection should be assumed where a service provider directs its activities to a Member State. It remains to be clarified precisely what “directing” activities means and if this is intended to go beyond “targeting”.

## A multi-layered approach

- 11 Consultations conducted prior to publication of the proposed DSA revealed concerns among stakeholders that a one size fits all approach to digital regulation would lead to a disproportionate burden being placed on smaller businesses and new market entrants. Taking this into account, the DSA adopts a multi-layered approach which imposes obligations determined by reference to four categories of intermediary service provider, with the obligations placed on a service provider building cumulatively if they fall into more than one category.
- 12 The four categories of service provider are:
- Intermediary Service – a basic level of obligations will apply to these services which cover:
    - a ‘mere conduit’ service that consists of the transmission of information provided by a user of the service, or the provision of access to a communication network;
    - a ‘caching’ service that consists of the transmission of information provided by a user of the service, involving the automatic, intermediate and temporary storage of that information, for the sole purpose of making the information's onward transmission more efficient upon another user's request; or



- a 'hosting' service which consists of the storage of information provided at the request of a user of the service.

Examples: Services offering network infrastructure, such as internet service providers or domain name registrars.

- **Hosting Service** – additional obligations apply to hosting services.

Examples: The hosting category will apply to those offering services such as cloud or web hosting services.

- **Online Platform** – further obligations still will apply to a provider of a hosting service which at the user's request stores and disseminates information to the public, unless that activity is a minor and purely ancillary feature of another service and for objective and technical reasons cannot be used without that other service.

Examples: Platforms such as online marketplaces, app stores and social media/content sharing platforms.

- **Very Large Online Platform ("VLOPs")** – the most burdensome obligations will apply to an Online Platform which provides services to an average 45 million plus active users per month. This figure is intended to reflect 10% of the population of the EU and can be increased or decreased to remain in line with this aim.

Examples: The 'big tech' companies that make up the largest online marketplaces, social media and content sharing platforms.

# What are the obligations?

- 13 The DSA imposes obligations that increase cumulatively depending on the type and extent of the activity conducted by the service provider. This means that providers who act only as mere conduits or merely cache information have the fewest obligations, with Hosting Services, Online Platforms and VLOPs each having increasing obligations.

## Over-arching provisions

### Modernisation of the intermediary liability regime

- 14 For intermediary service providers (those who cache or act as a mere conduit), there are only a limited number of new obligations introduced via the DSA. Chapter II of the draft Act, which deals with this category of provider, starts by restating the existing intermediary liability regime. Key principles from the e-Commerce Directive remain and are fully restated in the draft legislation, specifically the intermediary liability exemptions or “safe harbours” and the prohibition for Member States to introduce a “general monitoring” obligation on intermediaries with respect to the information that they transmit or store.
- 15 However, an important qualification has been added to the exemptions in Article 6, which is being referred to as the “Good Samaritan” provision. Article 6 makes clear that service providers won’t be ineligible for the liability exemptions “solely” because they take voluntary steps to detect, identify and remove or disable access to illegal content. This is intended to encourage platforms to take own-initiative steps to tackle illegal content without fear of recrimination. However, the DSA also makes clear that actual knowledge of illegal content may be gained via such voluntary activities and, if so, the duty to remove the content expeditiously, in order to benefit from the exemption, will still apply. This may prove a hostage to fortune for those who attempt to take more own-initiative steps.

### Illegal Content

- 16 The European Commission has stated that it does not purport to define what illegal content is, since this remains a matter for applicable national and EU law. However, the draft act refers to ‘Illegal Content’ broadly as a term that also covers information relating to illegal content, products, services and activities. Examples would include activity falling within a number of threat categories, including Pornography, Extremism, Graphic Violence, Drugs and Weapons. Recital 12 outlines some examples of activities deemed illegal under applicable law such as illegal hate speech or terrorist content and unlawful discriminatory content, or that relate to activities that are illegal, such as: the sharing of images depicting child sexual abuse, unlawful non-consensual sharing of private images, online stalking, the sale of non-compliant or counterfeit products, the non-authorised use of copyright protected material or activities involving infringements of consumer protection law.
- 17 It is notable that the DSA stops short of regulating legal but harmful content (for example, content relating to bullying or self-harm). An explanatory memorandum accompanying the DSA explains that the EU has taken this decision mindful of the delicate balance that needs to be struck between regulating this type of content and the implications such regulation would have on freedom of expression. In contrast, the UK’s draft Online Safety Bill, not only provides for a duty of care for companies towards users of their platform in relation to illegal content, but also proposes that a smaller group of companies deemed to pose a higher risk to users will also need to take action against legal but harmful content.

## Obligations that apply to all service providers

### Dealing with Take Down Orders

- 18 Service providers are obliged to act in accordance with any valid take down order received from a Member State's courts or administrative authority. The DSA stipulates certain information which must be contained in such an order, including: (a) why the information is illegal content; (b) information enabling the location or identification of the illegal content; and (c) information about redress available to the service provider. Service providers must confirm to the issuing authority, without undue delay, what action has been taken and when.

### Dealing with Disclosure Orders

- 19 Similarly, service providers must respond without undue delay to court/administrative orders requiring them to provide information about individual users of their services. Disclosure orders must contain reasons explaining why: (a) the information is required; (b) the requirement is necessary and proportionate to determine the user's compliance with applicable laws; and (c) details of redress available to the service provider and to the impacted user(s).

### Point of Contact

- 20 All service providers shall establish a single point of contact which allows for direct electronic communication with the authorities of Member States and the European Commission. This information shall be made public and service providers will also be obliged to make clear which languages can be used to communicate with them.

### Legal Representatives

- 21 Where a service provider is not established in an EU Member State, it must have appointed a legal representative in a Member State in which it provides services, to allow for effective oversight and, where necessary, enforcement of the DSA. Service providers will have to provide the name, address and contact details of their legal representative to the Digital Service Coordinator in the Member State where the legal representative is located (see the Competent Authorities section below for further information about the role of Digital Service Coordinators). This legal representative can also function as a point of contact. It is important to note that this legal representative can be held liable for non-compliance with obligations under the DSA.

### Terms and conditions

- 22 All service providers must have clear and unambiguous, publicly available and easily accessible terms of use for their service. These terms should include any restrictions they impose in relation to content provided by users. This will include information on policies, procedures, measures and tools used by the service provider for the purpose of content moderation, including any use of algorithmic decision-making and human review.

### Annual Reports

- 23 All service providers must produce annual transparency reports on the activities they undertake which are covered by the DSA. The annual reports of all service providers must cover information on: (a) actions they have taken at the request of national authorities, which will include the number of orders they have received and be categorised by the type of illegal content concerned; and (b) details of any content moderation activity they have undertaken on their own initiative.

## Obligations that apply to Hosting Services, Online Platforms and VLOPs

### User reporting function

- 24 Providers of Hosting Services must have mechanisms which enable users to report specific items of content which they believe contain illegal content. The DSA sets out that these reports will include:

(a) an explanation of why the user believes the information is illegal content; (b) the URL(s) where the information is located; (c) the name and email address of the user; and (d) a statement confirming the good faith belief of the user that the report is accurate and complete. If a report includes all of these elements, a service provider will be considered to have actual knowledge or awareness of the content for the purposes of considering the intermediary liability exemptions. Service providers will be obliged to respond to reports they receive “without undue delay”, setting out decisions they make and explaining how those decisions can be challenged.

### Statement of reasons

- 25 Hosting service providers must explain decisions they make to remove or disable access to specific content created or uploaded by users of their service and explain how the decisions they have made can be challenged. The reasons provided to the user must include: (a) whether the decision includes any removal or disabling of the content; (b) the facts that led to the decision being taken, including if they were taken following receipt of a report from another user; (c) whether any automated means were used in taking the decision; (d) if illegal content is involved, a reference to a legal ground relied upon and an explanation of why the information was considered to be illegal content; (e) if the information is deemed to violate the service provider’s terms and conditions, reference must be made to the relevant term relied upon; and (f) information about how the user can challenge the decision.

### Annual reports

- 26 A Hosting Service provider’s annual transparency report must additionally include information regarding notices they have received from members of the public about illegal content on their platform. This will be categorised by the type of illegal content concerned and will include details of any action taken following receipt of these notices, including whether action was taken on legal grounds or pursuant to the service provider’s terms and conditions. Information will also have to be provided as to the average time taken to act on such notifications.

## Obligations applying to Online Platforms and VLOPs

### Internal complaints process

- 27 Online Platforms must have a system for handling complaints about their decisions to remove or suspend content or a user account, on the basis that a user has published illegal content or has breached the service provider’s terms and conditions. The complaints process must be free of charge and available electronically. Service providers will be under obligations to ensure their complaints process is easy to use, carried out in a prompt and effective manner and that decisions will be reversed if the content is not found to be illegal or in breach of the provider’s terms and conditions. Online Platforms will have to inform users of any decisions taken following receipt of a complaint and provide information about how that decision can be challenged. Further, Online Platforms will be required to ensure that when a user is informed of its decision, this is not solely done on the basis of automated means.

### Out-of-court dispute settlement

- 28 Online Platforms will be required to engage with certified out-of-court dispute settlement bodies in relation to decisions they take to remove or suspend content or a user’s account. This settlement process will be accessible electronically. If a settlement body decides in favour of a user, the service provider will be required to bear the reasonable costs of the user. If the decision favours the service provider then each party shall bear its own costs. Users will not be obliged to use the out-of-court procedures and will remain free to pursue legal action in the courts as an alternative.

### Trusted flaggers

- 29 Certain entities, such as non-governmental organisations or the internet referral units of national enforcement authorities, will be able to apply to Member States to be recognised as ‘trusted flaggers’. To become a trusted flagger, an entity will need to demonstrate:

(a) it has the expertise and competence required to detect, identify and notify of illegal content; (b) it represents collective interests and is independent of any Online Platform; and (c) it carries out its activities for the purpose of submitting notices in a timely, diligent and objective manner. Online Platforms will be required to ensure that reports submitted by trusted flaggers are treated with priority and are decided upon without delay.

#### Abuse monitoring

- 30 Online Platforms are required to monitor whether certain users are misusing/abusing their platform by repeatedly submitting illegal content, or whether particular entities are repeatedly submitting unfounded complaints about illegal content. Where abuse is identified, Online Platforms will issue a warning to those concerned, following which they shall suspend their services to those concerned for a reasonable period of time. A service provider's policy for dealing with this type of abuse of their platform will need to be set out in a detailed manner in their terms and conditions.

#### Reporting suspicions of illegal activity

- 31 Any Online Platform that becomes aware of information giving rise to a suspicion that a serious criminal offence involving a threat to life or safety of persons has or is likely to take place, must report this to the relevant authorities. If the Online Platform cannot establish the relevant Member State in which they believe the activity has or will occur, they will inform the authorities in the Member State in which they are established (or where they have their legal representative) or inform the European Union Agency for Law Enforcement Cooperation (Europol).

#### Obtain, verify and publish information about traders

- 32 To improve the traceability of traders, Online Platforms who operate as online marketplaces must take steps to obtain and verify certain identity information from traders that use their services. The information to be obtained, verified and made available to consumers includes: (a) contact details; (b) the trade register in which the trader is registered; and (c) a self-certification by the trader committing to only offer products or services that comply with EU law. If information is suspected of being inaccurate and the trader refuses to correct that information, the trader must be suspended from the platform. Online Platforms must also design their interfaces in such a way as to allow traders to provide legally required pre-contractual information and product safety information.

#### Advertising

- 33 Online Platforms that display advertising must ensure that users can identify advertisement(s) and see on whose behalf it is being displayed. Users will also have rights to meaningful information about the main parameters which have been used to select them as a recipient for the advert.

#### Annual reports

- 34 An Online Platform must provide a report at least once a year that will include:
- Any use made of automatic means for the purpose of content moderation, including a specification which covers the purpose of the automated means and indicators of the accuracy of the automated means in fulfilling those purposes.
  - The number of complaints they have received from users via their internal complaints handling procedures, including the basis of those complaints and decisions taken as a result.
  - The time taken to deal with complaints and the number of instances in which the service provider's original decision was reversed following a complaint.
  - The number of disputes submitted to out-of-court settlement bodies along with their outcomes and the average time taken to complete the procedure.
  - The number of suspensions imposed as a result of abuse monitoring, including distinguishing between suspensions for the provision of illegal content and suspensions relating to unfounded complaints.

## Obligations that apply to Very Large Online Platforms only

### Identify 'systemic risks'

- 35 LOPs must conduct risk assessments to identify, analyse and assess any systemic risks associated with their services, including the risk of their services: (a) being used to disseminate illegal content; (b) having a negative effect on fundamental rights, including those to private and family life and freedom of expression; and (c) being intentionally manipulated in ways which negatively affect public life, including in relation to public health, minors, civil discourse, electoral processes and public security.

### Mitigate 'systemic risks'

- 36 In addition to identifying systemic risks, VLOPs will also have to take action to mitigate the systemic risks they identify. What mitigation action is required will depend on the circumstances, but might include: (a) adapting content moderation or recommender systems; (b) adapting the features or functioning of the service or their terms and conditions; (c) targeted measures to limit the display of certain advertising; (d) initiating or adjusting their cooperation with trusted flaggers; or (e) initiating or adjusting cooperation with other Online Platforms (or VLOPs) through codes of conduct and/or crisis protocols.

### Independent audit

- 37 Once a year, VLOPs will be subject to an independent audit at their own expense to assess their compliance with the EU DSA and any commitments they have made through codes of conduct and crisis protocols. These audits will be conducted by independent experts in risk management who will produce an audit report. If a negative audit report is received, VLOPs will be required to address issues identified in an audit implementation report within one month.

### Recommender systems

- 38 Where VLOPs use recommender systems, their terms and conditions must clearly set out: (a) the main parameters used by those systems; and (b) any options for users to modify or influence those parameters, including at least one option which is not based on profiling.

### Access to data

- 39 Following receipt of a reasoned request, VLOPs must provide the Digital Services Coordinator in the Member State in which they are established with electronic access to data to enable the Digital Services Coordinator to monitor their compliance with the EU DSA and to conduct research into the systemic risks faced by that particular VLOP. The DSA recognises that further delegated acts will be required to address concerns surrounding data protection, trade secrets and security arising out of this obligation.

### Compliance officers

- 40 VLOPs will be obliged to appoint compliance officers to: (a) cooperate with national Digital Service Coordinators and the Commission; (b) organise and supervise activities relating to independent audits; (c) train management and employees of VLOPs about their obligations under the DSA; and (d) monitor compliance with the DSA. The DSA makes clear that these compliance officers should report to the highest management level of the VLOP.

### Advertising

- 41 VLOPs must publish information about adverts they display, including on whose behalf they are displaying, whether they are targeted, the parameters used for targeting and the number of users to whom adverts are displayed.

### Annual reports

- 42 A VLOP's annual report will include information about the systemic risks they have identified and their mitigation measures for those systemic risks. It will also include information as to audits that have been undertaken and the implementation of any recommendations arising out of those audits.



# What are the penalties / enforcement measures?

## Competent authorities

43 There will be three primary competent authorities:

- **Digital Service Coordinators:** Each Member State will appoint its own Digital Services Coordinator. This will be a primary national authority responsible for supervising the intermediary services established in their Member State. The Digital Service Coordinators will have enforcement powers in respect of those service providers who are under the jurisdiction of their Member State, including the power to impose fines.
- **European Board for Digital Services (EBDS):** The Digital Services Coordinators will cooperate with the EBDS, a new independent advisory group at EU level. The EBDS will issue reports and recommendations and will coordinate joint investigations by Digital Service Coordinators. It is not intended to be seen as a regulator and is anticipated to have a similar role and remit to the European Data Protection Board.
- **European Commission:** The European Commission will be given new powers of supervision, investigation, monitoring and enforcement in respect of VLOPs.

## Sanctions

44 The sanctions mechanisms under the proposed DSA provide for enforcement at both Member State and EU level:

- **Member States:** Each Member State will specify penalties in their national laws which will be in line with the requirements set out in the EU DSA. The exact amount of the financial fines will be at the discretion of Member States but could be up to 6% of the service providers' annual income or turnover.
- **European Commission:** For VLOPs, the Commission will have direct powers of enforcement and the ability to impose significant fines of up to 6% of a VLOP's global income or turnover in the preceding year. The Commission will also have the power to impose a periodic penalty payment on VLOPs of up to 5% of their average daily turnover.

45 In addition to imposing monetary fines, the DSA provides for various other investigatory and enforcement measures, including:

- **Provision of information:** Digital Service Coordinators can require the provision of information about a suspected infringement, either from an intermediary service provider or a trader, within a reasonable period of time.
- **On-site inspections:** Digital Service Coordinators can conduct on-site inspections of any premises of an intermediary service provider or any other trader, to examine, seize or take or obtain copies of information relating to a suspected infringement.

- **Explanations:** Digital Service Coordinators can seek explanations from members of staff or representatives of a service provider in relation to suspected infringements and record their answers.
- **Interim Measures and Orders:** In cases where there is an urgency due to the risk of serious damage for the recipients of a service, the Commission may order interim measures. Digital Service Coordinators can also order the cessation of an infringement and impose proportionate remedies to bring the infringement to an effective end.

# What steps should companies be taking to comply?

- 46 The draft EU DSA legislation, which is currently subject to debate and amendment, provides some useful detail as to what measures are expected and how companies should implement them. However, more specific guidance will only be available when the law is set in stone and the regulatory bodies turn their attention to producing detailed codes of practice.
- 47 To the extent that your organization may be affected by the proposals, legal advice should be obtained as to the extent to which you will be affected by the obligations contained in the proposed DSA and what those obligations will require of you.
- 48 In the meantime, based on the content of the draft Act, as well as existing guidance and voluntary codes of practice, it is possible to suggest some steps which are likely to be considered good practice in relation to the various aspects of the proposed regime.

## Status under the DSA

- 49 One important task service providers need to undertake is to understand exactly where their services fall within the pyramid of obligations placed on service providers.
- 50 What changes are required to the way your organization operates is likely to be affected by whether it is deemed an Intermediary Service, Hosting Service, Online Platform or Very Large Online Platform under the proposals. It will therefore be vital that organizations establish at an early stage which category they will fall under.

## Use of automated content moderation

- 51 There is a focus on the use of automated means for content moderation, with obligations to disclose where such technology is used, including the use of algorithmic decision-making. Online Platforms must also report on the accuracy of such means and the safeguards applied to them. This may require service providers to reconsider their contractual arrangements with third party providers of automated content moderation.
- 52 Service providers ought to consider the advantages of automated content moderation, in light of their obligations to tackle illegal content under the DSA, including the accuracy and safeguards attached to any automated tools they currently make use of, as this will be disclosable under the DSA.

## Participate in the development of voluntary industry standards

- 53 The DSA makes clear that the European Commission encourages the development of voluntary industry standards.
- 54 While it is anticipated that these industry standards should be voluntary, it is expected they will become markers of best practice. A failure to adhere to any standards agreed on an industry-wide basis would therefore not reflect well on service providers in front of a future regulator.

- 55 Given that development of these concepts remains at a nascent stage, there is an opportunity for service providers to take an active role in the development of any voluntary codes of practice which they anticipate will be relevant to them and could benefit their business.

## Advertising contracts and practices

- 56 Service providers may need to consider whether the DSA will change what they require contractually from providers of advertising services. Online Platforms and VLOPs will face significant obligations to ensure that the beneficiaries of advertisements can be clearly identified by a user and to ensure that any targeting measures are identified and explained to users.
- 57 It would therefore be prudent for service providers to:
- engage with their providers of advertising services on these issues and ensure that appropriate contractual rights are in place; and
  - develop and publish a clear and accessible policy in relation to advertising practices.

## Updating terms and conditions

- 58 The DSA contains express obligations as to what information must be included in a service provider's terms and conditions and the level of detail those matters are required to be covered in.
- 59 Generally, it will be good practice for service providers to ensure that their terms and conditions are not overly long or complex and they should be easily understandable to their users. Terms should also be easily accessible, and users should be notified when significant changes to those terms are made, which may involve prompting users to read and acknowledge relevant sections.

## Proactive monitoring of decision-making processes and terms and conditions

- 60 Any service providers who fall within the VLOP category should develop processes by which they keep their decision-making processes and terms and conditions on content moderation under constant review.
- 61 As part of their obligations to mitigate systemic risks, VLOPs will be required to consider adapting their decision-making processes or their terms and conditions if necessary. A proactive approach to content regulation is likely to be viewed more favourably by any future regulator than a reactive one.

## User reporting functionality and complaints procedures

- 62 The DSA sets out in some detail the level of user reporting functionality a platform will be required to have and the complaints procedures they must ensure are in place to deal with grievances of users. Online Platforms should therefore carry out an assessment of their current procedures in order to ascertain any changes that might be required to bring them in line with the more onerous obligations set out in the DSA.
- 63 Another area service providers should be considering in the lead up to the DSA is the amount of resources they currently allocate to dealing with content moderation and user complaints and the extent this will need to be increased in order to comply with the requirements of the DSA. The DSA contains significant obligations to inform users of decisions taken by service providers and to provide regular reporting on the measures being taken by the service provider.

# Closing thoughts

We hope you have found this whitepaper useful and that it has provided some clarity on the legal obligations proposed within the EU Digital Services Act, whether your organization is in scope to comply with the new legislation and, if so, which of the four service provider tiers will apply to your services.

Regardless of their tier, all organizations in scope will have to comply with Take Down Orders relating to illegal content on their platforms. However, the sheer scale of user-generated content enabled by social media, smartphones and increased connectivity, have made it increasingly difficult for human moderators to stay on top of the task of removing illegal images and videos from their websites, online gaming communities, and digital platforms. Social media statistics tell their own story:

- 300M** Of the 300 million daily posts to Facebook, 35.73% are photographs and 15.09% are videos.
- 95M** 95 million photographs are uploaded to Instagram daily.
- 350K** 350,000 hours of video are streamed on Periscope every 24 hours.
- 500HRS** 500 hours of video are uploaded to YouTube every minute.
- 456K** There are 456,000 Tweets a minute.
- 950YRS** It has been estimated that it would take someone 950 years to watch all the Snaps shared daily on SnapChat
- 82%** By 2022 InterDigital/Futurescope anticipate that videos will make up 82% of consumer internet traffic.

A major challenge for organizations today is that a small percentage of these images and videos are deeply distressing and risk driving users off your site forever. When rogue actors abuse your digital service to upload images or live-streamed footage depicting or inciting graphic violence, child abuse and illegal pornography, this material pollutes online communities and spills over into real world harms.

Germany's Network Enforcement Act (NEA) already requires Social Network Providers that have more than 2 million registered users in Germany, to remove 'manifestly unlawful' content within 24 hours of receiving a complaint. Millions of daily posts can no longer be checked by humans alone. Delays in removing harmful content risk online user safety and penalties for non-compliance.

As you've just read, the EU Digital Services Act proposes that any digital platform that is used by more than 45 million members of the European population is deemed 'systemic' and therefore has a greater duty to oversee the content posted by users. Digital Services Coordinators will gain the power to directly sanction platform operators who fail to remove illegal content posted by traffickers, pornographers, extremists and rogue traders, and will be empowered to impose penalties of up to 6% of global turnover in the preceding year.

To keep illegal content off their platforms, the largest social media companies have outsourced their content moderation to contractors who employ many thousands of content moderators to review posts and flag the most toxic content for removal. However, even these huge, dedicated teams struggle to keep up with the volume of content being uploaded.

For smaller websites and fledgling online communities, managing content moderation as they scale up and attract more users within Europe has become a strategic issue for the C-suite.

If your organization is struggling to protect your online communities from harmful visual material so that you can create a positive user experience and comply with impending laws, we are here to help. Our award-winning AI-powered visual intelligence technology can stop illegal and harmful images, videos and live-streams from being uploaded to your site, leaving only the more nuanced visual content for human review.

To discuss your automated content moderation needs and discover why leading brands are using our technology, please visit [www.image-analyzer.com](http://www.image-analyzer.com) to request a demo, or contact me.

Sincerely,

**Crispin Pikes**  
CEO

Image Analyzer  
[crispin.pikes@image-analyzer.com](mailto:crispin.pikes@image-analyzer.com)

## IMAGE ANALYZER

Image Analyzer provides artificial intelligence-based content moderation technology for image, video and streaming media, including live-streamed footage uploaded by users.

Its technology helps organizations minimize their corporate legal risk exposure caused by employees or users abusing their digital platform access to share illegal or harmful visual material. Image Analyzer's technology has been designed to identify visual risks in milliseconds, including illegal content, and images and videos that are deemed harmful to users, especially children and vulnerable adults.

Founded in 2005, Image Analyzer holds various patents across multiple countries under the Patent Co-operation Treaty.

The company is a member of the Online Safety Tech Industry Association (OSTIA).

---

## Bird & Bird

[twobirds.com](http://twobirds.com)

Abu Dhabi & Amsterdam & Beijing & Bratislava & Brussels & Budapest & Copenhagen & Dubai & Dusseldorf & Frankfurt & The Hague & Hamburg & Helsinki & Hong Kong & London & Luxembourg & Lyon & Madrid & Milan & Munich & Paris & Prague & Rome & San Francisco & Shanghai & Singapore & Stockholm & Sydney & Warsaw & Satellite Office: Casablanca

The information given in this document concerning technical legal or professional subject matter is for guidance only and does not constitute legal or professional advice. Always consult a suitably qualified lawyer on any specific legal problem or matter. Bird & Bird assumes no responsibility for such information contained in this document and disclaims all liability in respect of such information.

Bird & Bird is an international legal practice comprising Bird & Bird LLP and its affiliated and associated businesses.

Bird & Bird LLP is a limited liability partnership, registered in England and Wales with registered number OC340318 and is authorised and regulated by the Solicitors Regulation Authority. Its registered office and principal place of business is at 12 New Fetter Lane, London EC4A 1JP. A list of members of Bird & Bird LLP and of any non-members who are designated as partners, and of their respective professional qualifications, is open to inspection at that address.

49377090.1