

Online Harms

A comparative analysis



Online Harms: A comparative analysis

Foreword

The evolving regulatory response to online harms

03

Thematic review

1. Who is regulated?

05

2. What types of harm do platforms need to protect their users from?

07

3. Is the focus on individual pieces of content or systems and processes?

09

4. What do platforms have to do to comply?

10

5. What are the consequences if a platform fails to comply?

13

6. What happens next?

15

Country by country review

Australia

18

European Union

21

France

24

Germany

26

Ireland

28

Singapore

30

United Kingdom

32

United States

34

The evolving regulatory response to online harms

Foreword: The potential for harm

For the first three decades of its existence, much of the public discussion around the internet focused on the ways in which the technology had improved our lives. Information that was previously buried in encyclopaedias became accessible at the click of a button, interactive maps of the entire world were made available in the palm of your hand and it became possible to connect and converse with people all over the world.

Yet, as well as leading to these incredible enhancements to our lives, it has become increasingly clear that content on the internet can cause real harm too. Posts promoting extremism have been linked to terrorism, campaigns of disinformation and “fake news” have dogged democratic elections, and charities and governments have drawn attention to the horrifying volume of child sexual abuse imagery circulated online.

The risks of these “online harms” came into even sharper relief in 2020, as national lockdowns during the Covid-19 pandemic meant many of us spent more time online than ever before.

Regulatory proposals

A consensus has emerged in recent years that more needs to be done to combat “online harms”. As well as governments, charities and consumers calling for action, many of the large tech platforms have publicly called for regulation.

This has led to governments across the globe looking to replace the current patchwork of discrete laws and voluntary initiatives with more holistic regulation. In late 2020, ambitious proposals were announced by the EU, the UK and Ireland seeking to impose far greater obligations on organisations to tackle online harms. These proposals follow reforms in Australia, France, Germany and Singapore. This new wave of regulation has yet to crystallise in the U.S. where reform is the subject of much debate.

Country by country review

Though the list of countries looking to legislate in this area is ever-growing, in this publication we focus on the current legal position in Australia, France, Germany, Singapore and the U.S. and we look ahead to the proposals in the EU, Ireland and the UK. While there are some commonalities in approach, there is no consensus across these jurisdictions on how to regulate online content. This lack of harmonisation will undoubtedly create challenges for the largest online platforms preparing for regulation in multiple jurisdictions. We highlight the key aspects organisations will need to consider.

The regulatory balancing act – protecting against harm versus protecting fundamental rights

Regulating harmful content online is not a straightforward task. The global nature of the internet, with platforms and their users spread across nearly every jurisdiction in the world, huge volumes of content posted every second and the complex and disparate nature of the different types of harm, makes it difficult for governments to formulate a regulatory framework that is both effective and proportionate.

Governments and platforms have to balance their shared objective of reducing the amount of harmful content online with respect for users’ fundamental rights, such as the right to freedom of expression, and a range of other policy interests, from encouraging more competition to protecting data privacy.

Focus on hosting responsibilities

There are a wide range of laws that affect what can be posted by users online: such as laws relating to defamatory content, intellectual property rights and advertisements. This publication focuses not on what users can and can’t do online, but on the requirements, current and proposed, imposed on platforms which host content or facilitate contact between users – sometimes referred to as “intermediary liability”.

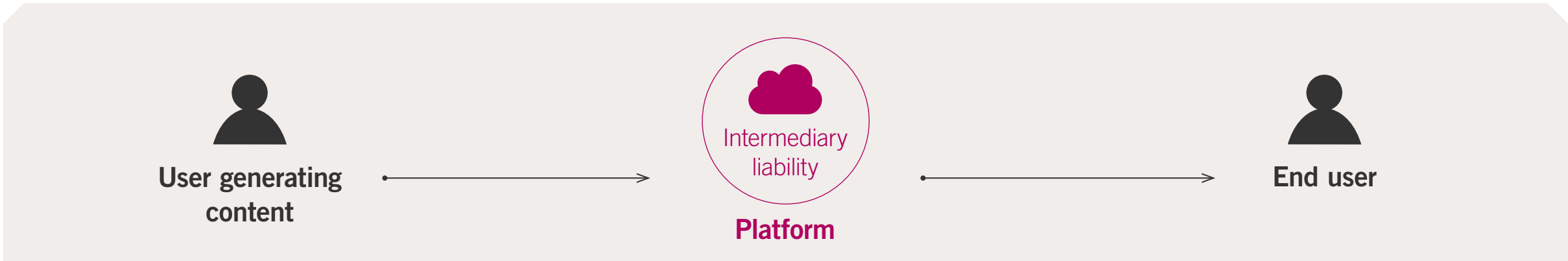
Not the end but the start of the journey

This publication covers rules that are already in force and new regimes that have been announced. However, this is merely a snapshot of a point in time. In the coming years, it is inevitable that the proposals will be tweaked, new rules will be introduced and entirely new areas of risk will emerge. As experience in other sectors has shown, regulation is rarely “done”. Rather, it has to adapt and change with the times. This is even more likely to be the case for a sector as innovative and fast-moving as tech. It’s therefore highly unlikely that the proposals outlined in the publication represent the end of the journey towards a safer internet, but rather the start.

We hope you enjoy the publication and would welcome your feedback. Please do get in touch with me or any of the team if you would like to discuss further.




Ben Packer
Partner, London
+44 20 7456 2774
ben.packer@linklaters.com





Thematic review




A comparative analysis in the form of a **thematic review** addressing the following key questions:

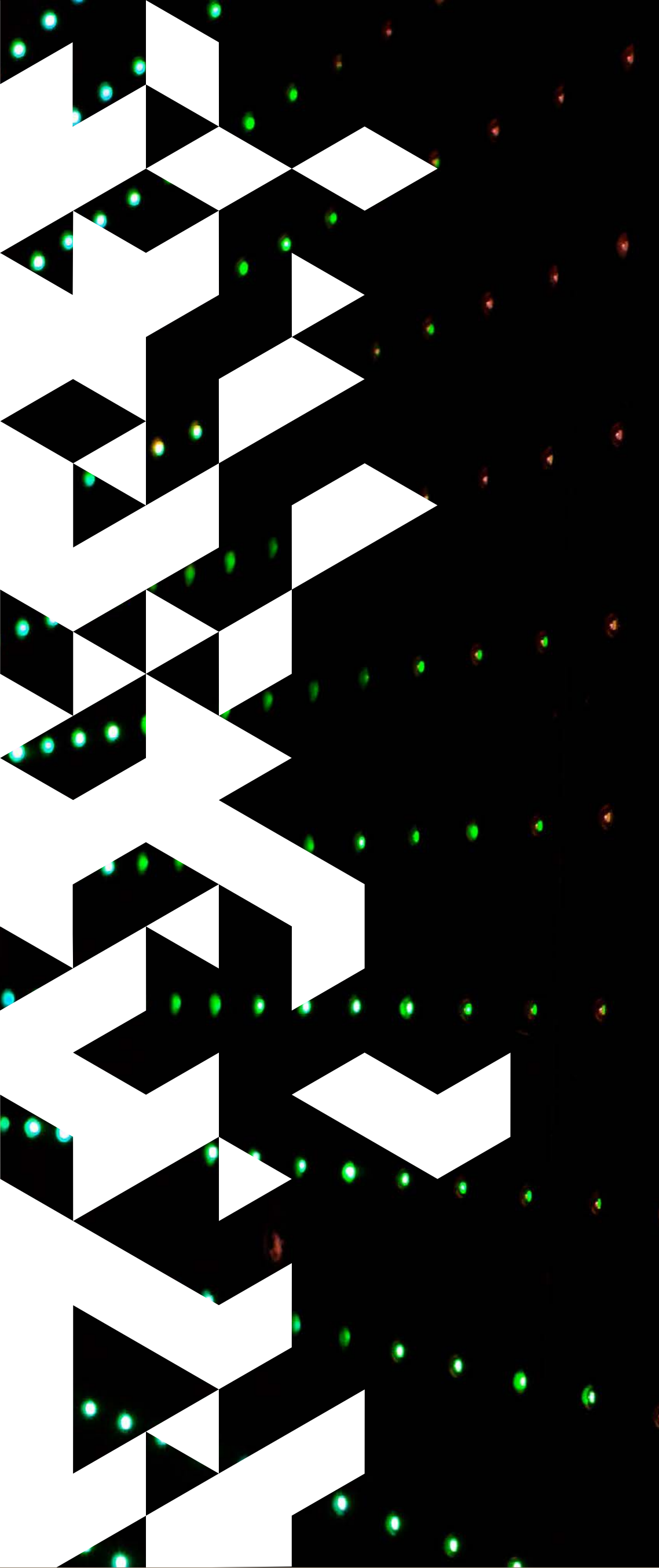
1. Who is regulated?
 2. What types of harm do platforms need to protect their users from?
 3. Is the focus on individual pieces of content or systems and processes?
 4. What do platforms have to do to comply?
 5. What are the consequences for failure to comply?
 6. What happens next?
- 

1. Who is regulated?

The question of which companies are in scope for each regime can be complex. While the respective legal definitions do not lend themselves to easy or precise comparison, set out below are the types of services which are likely to be caught under each.

Types of services

Services in scope	Social media platforms	Cloud hosting	Video content sharing	Video games with user interaction	Online marketplaces	Search engines	Private or user-to-user interactions
Australia (BSA, AVMA and EOSA)	✓	✓	✓	✓	✓	✓	✓
EU (DSA) 	✓	✓	✓	✓	✓	✓	?
France (LCEN and Avia Law)	✓	✓	✓	✓	✓	✓	✗
Germany (NetzDG)	✓	✓	✓	✗	✗	✗	✗
Ireland (Online Safety and Media Regulation Bill) 	✓	✓	✓	✓	✓	✓	✓
Singapore (Internet Code of Practice and POFMA)	✓	✓	✓	✓	✓	✓	✓
United Kingdom (Online Safety proposals) 	✓	✓	✓	✓	✓	✓	✓
United States (Section 230)	✓	✓	✓	✓	✓	✓	✗



Larger platforms

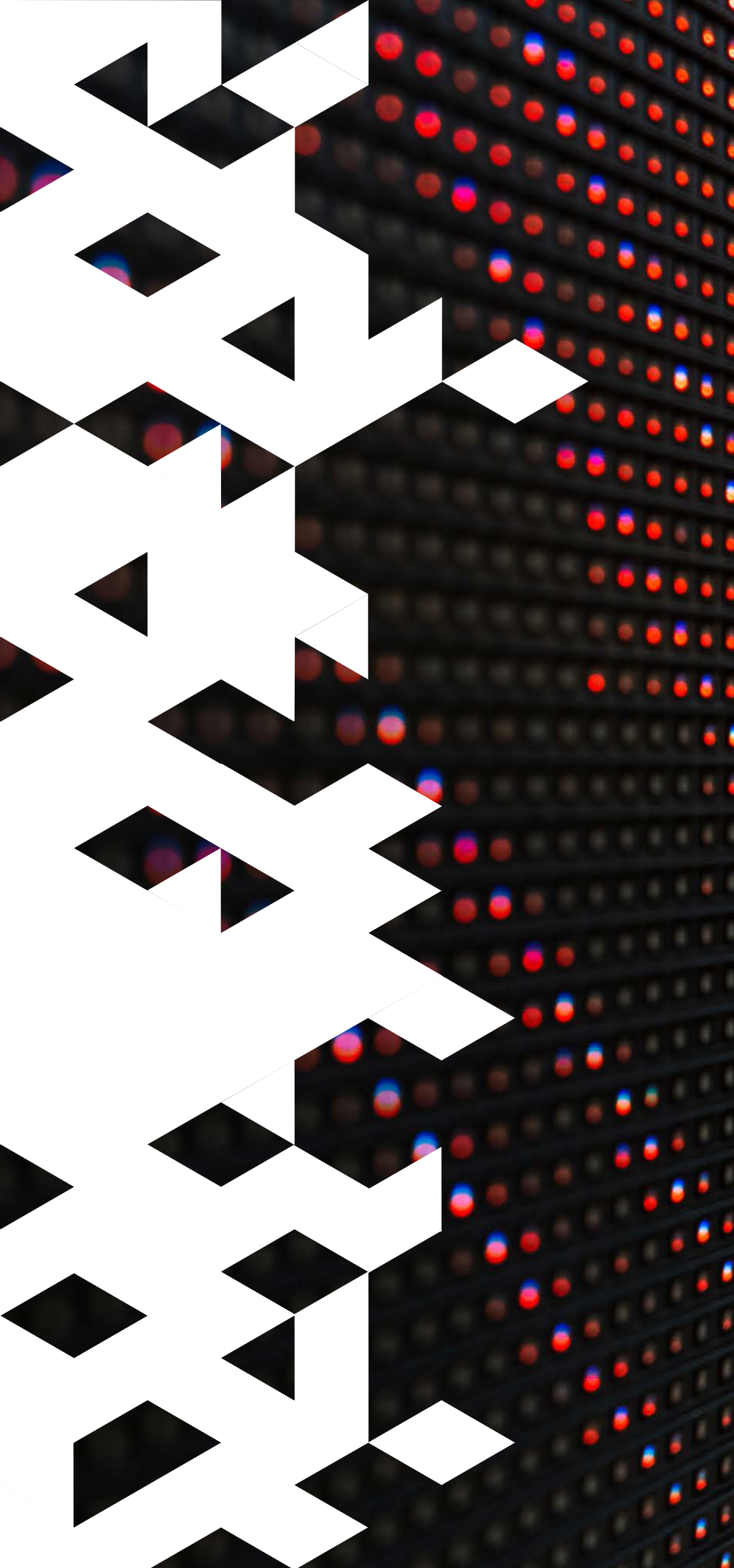
Both the EU and UK proposals impose more stringent obligations on larger platforms.

The obligations imposed by the EU proposals are scaled, depending on the EU’s judgement of the “*role, size and impact in the online ecosystem*” of different types of service. “Intermediary services”, those that offer network infrastructure, will be subject to the least stringent requirements, followed by “hosting services”, such as cloud and webhosting services, then “online platforms” such as online marketplaces, app stores and social media platforms and, finally, the most stringent obligations will be imposed on “very large online platforms”. “Very large online platforms”, or “VLOPs”, are those platforms with more than 45 million monthly active users in the EU.

In a similar vein, the UK proposals use the terminology of “Category 1 companies”, being those in-scope services which are considered to be “*high-risk and high-reach*”.

We have outlined the additional requirements VLOPs and Category 1 companies have to comply with in section 4, **What do platforms have to do to comply?** below.





2. What types of harm do platforms need to protect their users from?

A key focus of each of the regulatory regimes is to protect users from harmful content online. But what type of harmful content should platforms be expected to protect users against?

Each of the jurisdictions we looked at has tackled this question in a different way. All of the jurisdictions covered in this publication require platforms to take steps in relation to illegal content, such as terrorism-related content or child sexual abuse material. However, only some jurisdictions go beyond this and propose to regulate content that is “lawful but harmful”, such as disinformation (the EU proposals, Singapore and the UK) or the promotion of eating disorders (Singapore, Ireland and the UK). Even then, what is considered to fall within this category varies significantly from jurisdiction to jurisdiction. The UK and Irish proposals and the Singaporean regime seek to cover the broadest category of harms.

The table on the next page shows a high-level comparison of the types of content which are regulated in each jurisdiction. The underlying legal definitions will be specific to each jurisdiction.

Potential forms of harm:

- > Terrorism-related content
- > Child sexual exploitation and abuse content
- > Hate speech
- > Sale of illegal drugs and weapons
- > Disinformation
- > Violent content
- > Sexually explicit content involving adults
- > Cyberbullying
- > Promotion of eating disorders
- > Promotion of self-harm or suicide

“However, only some jurisdictions go beyond this and propose to regulate content that is “lawful but harmful”, such as disinformation (the EU proposals, Singapore and the UK) or the promotion of eating disorders (Singapore, Ireland and the UK).”

2. What types of harm do platforms need to protect their users from?

Jurisdiction and law	Do platforms' obligations extend beyond the removal of illegal content?	Status	Type of harmful content									
			Terrorism-related content	Child sexual exploitation and abuse content	Hate speech	Sale of illegal drugs and weapons	Disinformation	Violent content	Sexually explicit content involving adults	Cyberbullying	Promotion of eating disorders	Promotion of self-harm or suicide
Australia (BSA, AVMA and EOSA)												
EU (DSA)												
France (LCEN and Avia Law)												
Germany (NetzDG)												
Ireland (Online Safety and Media Regulation Bill)												
Singapore (Internet Code of Practice and POFMA)												
United Kingdom (Online Safety proposals)												
United States (Section 230)	N/A											

Key Is regulated or proposed to be Not regulated In force Proposal Possibly Roll over for further information




3. Is the focus on individual pieces of content or systems and processes?

The regimes can broadly be divided into those that focus on individual pieces of content and those that instead focus on the systems and processes that platforms must have in place:

- > **Focus on individual pieces of content:** The German, Australian and Singaporean regimes all do the former and impose rules relating to individual pieces of content. In practice, this means obligations to take down or disable access to individual pieces of content quickly (often known as “ex post” – or after the event – obligations).
- > **Focus on systems and processes:** In contrast, the EU (for online platforms and VLOPs), Irish and UK obligations relate to the overall systems and processes that a platform must have in place. This type of regime requires platforms to assess the risks of certain types of harmful content being present on the platform and take appropriate steps to mitigate those risks (sometimes called “ex ante” – or before the event – regulation).

The former approach has the advantage of imposing simpler obligations, where compliance (or the lack thereof) is easier to determine. However, advocates of a systems and processes approach argue that ex ante regulation has proved effective in other sectors and may mean that instances of harm are less likely to occur in the first place because the focus is on preventing harm, rather than merely responding after the event.

“However, advocates of a systems and processes approach argue that ex ante regulation has proved effective in other sectors and... instances of harm are less likely to occur in the first place because the focus is on preventing harm, rather than merely responding after the event.”



4. What do platforms have to do to comply?

Exactly what platforms have to do to comply with the respective regimes varies from jurisdiction to jurisdiction.

While there are fundamental differences in how the overarching obligations are framed, there are some relatively common obligations imposed on platforms across the various regimes. These include:



Removal/blocking of content following notification

All of the jurisdictions impose an obligation on platforms to remove or block certain types of content once it has been reported to them. The speed at which platforms must remove or block the content varies. The point at which the clock starts ticking also varies - in some jurisdictions a designated authority must have made the report, in others a report from a user is sufficient to start the clock.

- > **Germany:** The strictest regime is the NetzDG in Germany. “Manifestly” unlawful content must be removed or blocked within 24 hours of the platform receiving a report from a user or relevant organisation. The German Federal Office of Justice can impose a regulatory fine if a platform fails to comply.
- > **Australia:** By way of contrast, in Australia, under the BSA a platform must remove or block content as soon as practicable, but only after receiving a notification from the Australian eSafety Commissioner. Under the AVMA, a platform must expeditiously remove abhorrent material. An offence is established where the platform was reckless (i.e. they were aware of a substantial risk that their platform could be used to access the material and they did not remove it). This means that there is also an implied requirement to remove content following a user report. Under the EOSA, a platform must remove cyberbullying material within 48 hours of receiving a user request to remove the material. A platform must also remove intimate images which have been shared without an individual’s consent within 48 hours of receiving a removal notice from the eSafety Commissioner. Australia has recently published proposals to consolidate its online harms regimes and reduce the takedown time in removal notices from 48 hours to 24 hours for cyberbullying and intimate images which have been shared without an individual’s consent.

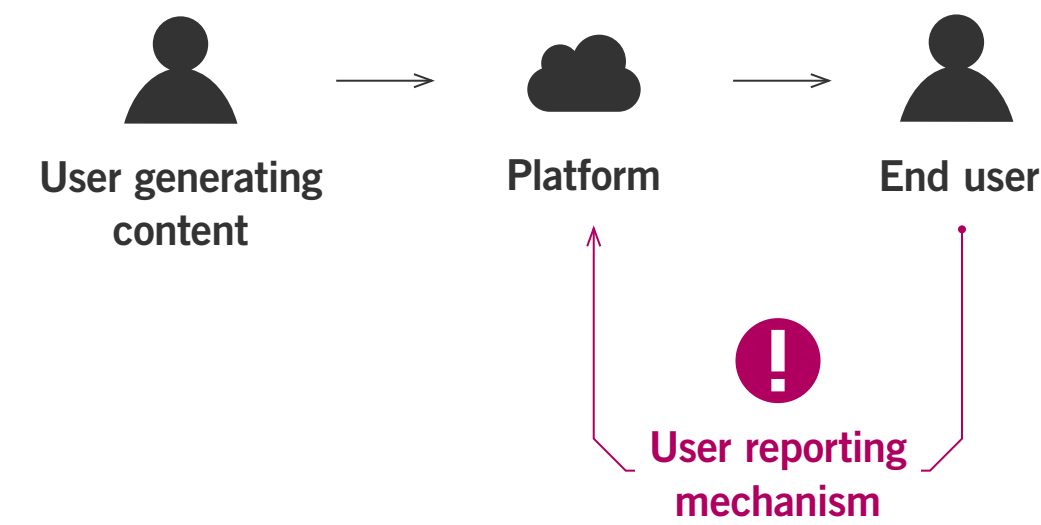
“While there are fundamental differences in how the overarching obligations are framed, there are some relatively common obligations imposed on platforms across the various regimes.”



Requirement to have a user reporting mechanism

Many of the jurisdictions require platforms to have a system in place to allow for user reporting of harmful, prohibited or illegal content (as defined in that jurisdiction). The regimes in the EU, France, Germany, Ireland and Australia all include a requirement to facilitate user reporting.

- > **Germany:** Platforms must provide a user reporting mechanism for illegal content, and it must be “effective and transparent” and “easily accessible”.
- > **EU:** The proposals include a similar requirement for providers of hosting services to provide an easy-to-access and user-friendly mechanism for users to report illegal content.
- > **France:** Similarly requires a system of user reporting to be provided for several criminal offences, including hate speech, child sexual abuse material, incitation of minors to engage in harmful games, human trafficking, procuring of terrorist content and illegal gambling activities.
- > **Australia:** Platforms must provide a user reporting mechanism for cyberbullying material.



Reporting to the authorities

Many of the jurisdictions impose obligations on platforms to report information to authorities about content found on the platform. However, the breadth of those obligations varies greatly.

- > **France:** Platforms are required to inform the competent public authorities of any content reported to them that may constitute one of a number of criminal offences, including hate speech, and to provide data which would help the authorities to identify the user who posted the content.
- > **Germany:** A similar obligation will apply in Germany from 1 February 2022, following a recent amendment to the NetzDG. Platforms will be required to report certain types of unlawful content to the Federal Criminal Police Office, such as online threats and hate crime content.
- > **EU:** The EU proposals provide that, if an online platform becomes aware of any information giving rise to a suspicion that a serious criminal offence involving a threat to life or to the safety of persons has taken place, is taking place, or is likely to take place, law enforcement or judicial authorities of the Member State concerned must be informed.
- > **Australia:** The obligation is not as broad. It only arises when a platform becomes aware that their service can be used to access abhorrent violent material, or that recording or streaming of such material has occurred or is occurring in Australia.



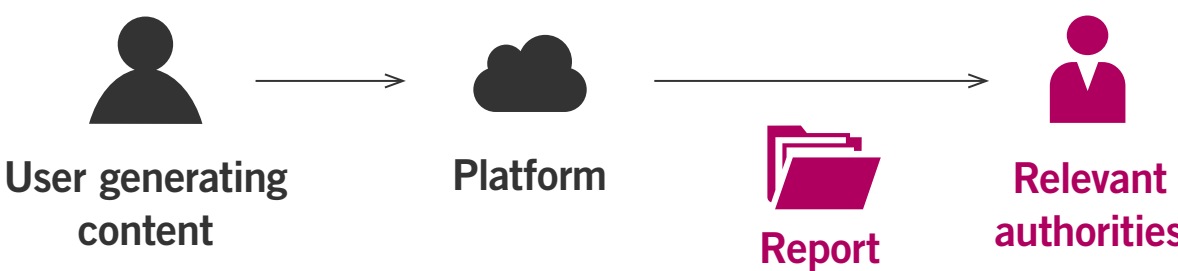
Extra obligations for the largest online platforms: going beyond illegal material

Perhaps unsurprisingly, many of the regulatory regimes expect more of the largest online platforms: as their user-base and extensive reach is seen as amplifying the risk of their services causing harm.

Under both the EU and UK proposals, the key additional obligation for the largest platforms is to conduct a holistic risk assessment that considers more than just risks arising from illegal content. The UK proposals require Category 1 companies to complete regular risk assessments to identify legal but harmful materials on their services and to take steps to mitigate those risks. Similarly, the EU proposals require VLOPs to undertake an assessment of any significant systemic risks their services pose. “Systemic risks” can relate to: the dissemination of illegal content; negative impacts on the exercise of fundamental rights (for example, the right to freedom of expression) or intentional manipulation of the service with an actual or foreseeable effect on public health, minors, civic discourse, the electoral process or public scrutiny. VLOPs will have to put in place reasonable, proportionate and effective measures to mitigate these risks. In addition, VLOPs have to arrange and pay for an annual external and independent audit of their risk management systems and processes.

- > **UK:** The UK government is also introducing targeted reporting requirements. It has said that it expects companies to report terrorist content to law enforcement where there is a threat to life or risk of an imminent attack. It has also said that it is “minded” to introduce a requirement for companies to report child sexual exploitation and abuse identified on the service.

The timescales for reporting also differ between jurisdictions. In France, the report must be made “promptly”, which has no exact definition, but the case law suggests that a delay of five days is too long. The EU proposals and the NetzDG in Germany (applicable from 1 February 2022) also require the report to be made “promptly”. In Australia, the report has to be made within a “reasonable” time. “Reasonableness” in this context is determined by the type and volume of the harmful content and the resources available to the provider. The UK government has not yet provided an indication of how quickly platforms will be expected to make reports.





Transparency reporting

Transparency reporting involves the publication of regular reports by platforms with the intention of providing external parties with insight into how platforms have dealt with certain issues over a period of time.

Many of the biggest online platforms already publish detailed transparency reports voluntarily. Typically, these cover a range of topics, including details of content moderation policies and decisions and reporting on how the platform interacted with law enforcement.

However, while platforms have broadly had choice over what to publish up to now, the new regimes will be more prescriptive.

Platforms which are subject to transparency reporting requirements in several jurisdictions will need to grapple with competing requirements on what must be published to satisfy their obligations in each jurisdiction without falling foul of any restrictions on what can be said. For example, U.S. law places restrictions on what can be said in transparency reports about requests made by authorities pursuant to the U.S. Foreign Intelligence Surveillance Act.

Platforms will need to work out whether they can meet all of the requirements in one report, or whether they need to produce tailored reports for certain jurisdictions.

- > **Germany:** All regulated bodies that receive over 100 complaints about illegal content per year must submit biannual reports to the regulator and publish them online. There are more reporting obligations in the pipeline.
- > **UK:** Category 1 companies will be required to publish annual transparency reports. The UK regulator will determine which categories of information platforms have to provide, with a view to keeping the requirements proportionate.
- > **EU:** The DSA will also impose significant transparency obligations on all in-scope companies, with the most onerous obligations falling on VLOPs. Not only will these platforms be required to publish the transparency reports required of all intermediary services, but they will be subject to additional obligations to publish their transparency reports more frequently (every six months), as well as publishing additional information such as their risk assessments and mitigation measures and details of the audit mentioned above.



Designated authorised person

Some of the regulatory regimes require platforms to name an authorised person to take responsibility for ensuring effective compliance.

- > **Germany:** All regulated entities must name an authorised person to receive service of regulatory and civil proceedings, and this person must be identified on the regulated entity’s website.
- > **EU:** The proposals include a requirement for VLOPs to appoint a compliance officer for the purposes of the regime. However, neither the EU proposals nor the German rules impose liability on the individual for non-compliance.
- > **UK:** The UK government has reserved its right to introduce criminal sanctions for senior managers into the UK regime. However, it has limited the scope of any such liability, which could arise only in situations where senior managers fail to respond “*fully, accurately and in a timely manner*” to information requests from Ofcom.
- > **Ireland:** The Irish government also plans to introduce criminal liability for senior managers in respect of specified offences committed by a designated online service where it is proven that the offence was committed with the consent or connivance of senior management, or where they have been acting with wilful neglect. Currently, the only specified offence for which senior management may be in scope for criminal liability is non-compliance by a regulated entity with a warning notice issued by the Media Commission requiring compliance.



What don’t platforms have to do?

A number of measures which have been suggested in the wider online harms debate have not been implemented in any of the regulatory regimes reviewed for this note.

Pre-moderation

For example, pre-moderation is often floated as a way to prevent harmful content from being uploaded in the first place. While this may work in specific instances (for example, the comments section of an online article), it is unworkable for the largest online platforms, like social media sites. Indeed, under the DSA, Member States will continue to be prohibited from imposing a general obligation to monitor on intermediary services.

Mandatory real world identification

Another measure often touted is the use of mandatory real-world identification of users. Individuals would have to provide some form of ID (such as a driving licence or passport) to register as users of platforms, and would only be able to post under their legal name.

While this measure has been introduced in very limited circumstances (such as for business retailers under the DSA), concerns have been raised about its wider application. It is estimated that 1.1 billion people worldwide do not have any form of officially recognised ID document.

A policy of requiring real world identification of users could create barriers to users based on their immigration status, gender identity or other sensitive factors.





5. What are the consequences if a platform fails to comply?

Credible and proportionate sanctions

Clearly, for any regulatory regime to be effective, it needs to incorporate credible and proportionate sanctions. As can be seen in the table overleaf, all of the proposed and current regimes except in the U.S. will allow authorities to levy financial penalties on platforms that have failed to comply with their obligations.

Some go further and allow for the corporate to be criminally liable for non-compliance with a regulator's requests (e.g. Ireland) or even for individual employees to be fined or imprisoned.

A tool of last resort

With online platforms typically offering services across borders, clearly there is a risk that regulators will struggle to enforce financial or criminal sanctions against platforms or their management. For that reason, several of the regimes allow regulators to take the “nuclear” option of blocking access to the platform altogether in their jurisdiction. For all of the frameworks that contain this option, this is positioned as a tool of last resort.

“With online platforms typically offering services across borders, clearly there is a risk that regulators will struggle to enforce financial or criminal sanctions against platforms or their management.”



5. What are the consequences if a platform fails to comply?

	Potential maximum fine for the platform	Possibility of corporate criminal liability?	Possibility of liability for individual directors or employees?	ISP blocking?	Other enforcement tools?
Australia	AUD 11.1 million (c. GBP 6 million) or up to 10% of annual global turnover, whichever is higher, per failure to expeditiously remove content	✓	✓	✓	<div>> Remedial directions and warnings</div> <div>> Court order that the person cease supplying an internet carriage service or providing the designating hosting service</div>
EU <div></div>	Up to 6% of global annual turnover, where a provider has been found to breach its obligations For VLOPs, periodic penalty payments up to 5% of the average daily turnover in the preceding financial year per day	✗	✗	✗	<div>> Requiring commitments from platforms that they will make their services compliant</div> <div>> Temporarily restricting access to the platform’s services</div> <div>> Periodic penalty payments of up to 5% of the average daily turnover of the platform</div>
France	EUR 1.25 million (c. GBP 1 million)	✓	✓	✓	<div>> Prohibition preventing the platform from carrying out its activities for up to five years</div>
Germany	EUR 50 million (c. GBP 43 million), depending on the seriousness of the non-compliance and taking into account the platform’s economic position	✗	✓	✗	✗
Ireland <div></div>	Administrative fines of up to EUR 20 million (c. GBP 17 million) or 10% of relevant turnover (whichever is higher) for the preceding financial year Criminal sanctions range from fines of EUR 5,000 (c. GBP 4,300) and/or up to 12 months’ imprisonment on summary conviction (depending on the nature of the offence). Sanctions for conviction on indictment have not yet been specified	✓	✓	✓	<div>> Information requests, investigations and audits</div> <div>> Issue directions through compliance notices and warning notices</div> <div>> Orders compelling compliance with warning notices or blocking access to the relevant entity</div>
Singapore	SGD 1 million (c. GBP 538,000) per non-compliance with a ministerial direction	✓	✓	✓	<div>> Directions to take down, disable or correct content</div>
United Kingdom <div></div>	GBP 18 million or up to 10% of global annual turnover, whichever is higher, for “failure to comply” with regulatory obligations	✗	✓	✓	<div>> Compel third parties to withdraw key services that make it less commercially viable for the company to operate within the jurisdiction</div>
United States	None (online platforms are broadly immune from liability for content generated by their users under section 230)	✓	✗	✗	✗

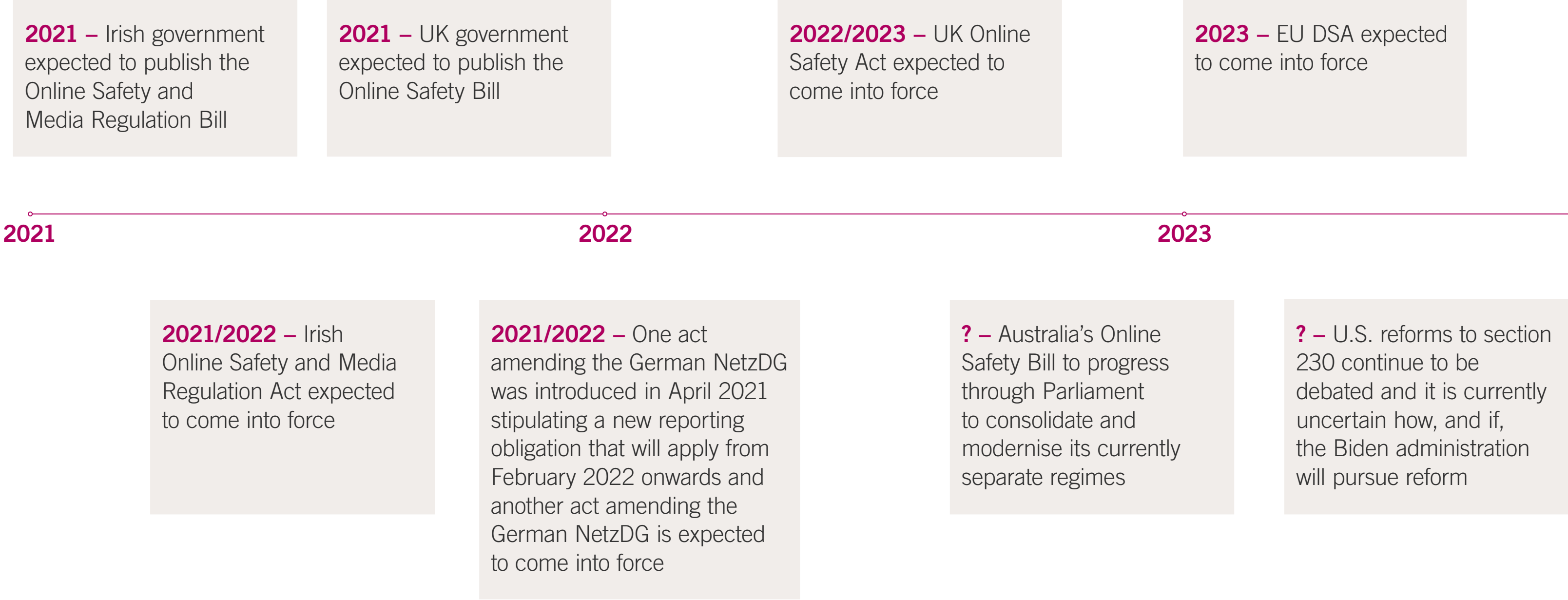
 Roll over for further information



6. What happens next?

Timeline for new legislation

The next few years will see the introduction of a number of new laws regulating online harms. For some of these, governments have given an indication of timing; for others, we are awaiting more concrete details of what happens next.



Broader trends

Though several of these regimes are already in force, the broadest and most ambitious ones are still in development and will inevitably change in some ways as they go through the legislative process. In the meantime, there are several broader trends that look set to play out in relation to online content.

> **Private action filling the void while the law is developed:** Platforms are facing unprecedented scrutiny from governments and the public. Put simply, platforms, users and society more broadly have recognised that the need for action on harmful content is so immediate that they cannot wait for regulatory regimes to be formulated and embedded.

While legislators formulate and progress their proposals, platforms have been forced to act. This has taken a range of forms: from Facebook setting up its “Oversight Board” for content moderation appeals, to Visa and Mastercard withdrawing payment services from PornHub due to concerns about unlawful content on the site; from Twitter and Facebook ultimately banning former President Trump from their services, to Amazon Web Services refusing to host Parler on its servers.

> **Potential clashes with other regulatory agendas:** The reduction of harmful online content is a priority in many of the jurisdictions that we cover in this publication. However, many of the same jurisdictions also pursue other policy goals: from encouraging more competition in online services to seeking to protect personal privacy. Many of these policy goals will inevitably come into conflict in due course. For instance, will stringent requirements around online content operate as a barrier to entry for new competitors? Does the expectation that platforms will examine more of their users’ content, including private messages in some jurisdictions, clash with expectations around personal privacy?

Ultimately, some countries may decide that one regulator which has to balance all of these competing interests when regulating the sector would be preferable to separate regulators all pursuing their own mandate.

None of the jurisdictions we looked at for this publication have yet taken this approach. However, in the UK, regulators are clearly conscious of the potential for clashes. The UK competition, data protection, communications and financial services regulators have established the “Digital Regulation Cooperation Forum”, with the aim of co-ordinating their regulatory approach across digital and online services.

> **Navigating divergent obligations:** As this publication demonstrates, the frameworks and obligations which have been or are being implemented differ greatly across jurisdictions, and we have only looked at a small sample.

Platforms will need to navigate numerous complex schemes and build and maintain relationships with new and additional regulators while offering a coherent and broadly uniform user-experience for their user-base globally.

> **The role of education:** Regulation alone cannot solve the problem of harmful content online. There is a real need to educate users on how to stay safe online, and to be sceptical and discerning about the content they are exposed to.

The UK regulator, Ofcom, has a statutory duty to promote media literacy. The UK Department of Education has also brought in new national standards for essential digital skills, including “Being safe and responsible online”.

> **Reduction in harmful content online ahead of new regulation:** The ultimate goal of all the regulation is to reduce the amount of harmful content online and to reduce the public’s exposure to that content. However, the data shows that online platforms are already making great strides ahead of the most ambitious regimes coming into force.

By way of example, the fifth evaluation of the EU’s voluntary Code of Conduct on Countering Illegal Hate Speech online found that, on average, 90% of the notifications of illegal hate speech were reviewed within 24 hours and 71% of the content was removed. Facebook has reported that in Q4 2020, it acted on 28 million pieces of content and that 98.1% of that violating content was found and flagged before users reported it.

Though there is no doubt that there is still some distance to travel on the journey to making the internet a safer place, it is clear that huge progress has already been made.





Country by country review





Australia



What framework is in place to regulate online harms?

Online harms are regulated by three laws:

- > the Broadcasting Services Act 1992 (the “**BSA**”). While the BSA has broad application, Schedules 5 and 7 apply specifically to internet content hosted outside Australia, or content hosted in or provided from Australia;
- > the Criminal Code Act 1995, as amended by the Criminal Code Amendment (Sharing of Abhorrent Violent Material) Act 2019 (the “**AVMA**”). The AVMA applies to material hosted on a “carriage service” (which includes telephone or internet services) which is provided within or outside Australia; and
- > the Enhancing Online Safety Act 2015 (the “**EOSA**”). The EOSA’s cyberbullying regime applies to cyberbullying material generally and cyberbullying material relating to an Australian child provided or posted on a social media service. The EOSA’s intimate images regime applies where the depicted person or user is ordinarily resident in Australia or the content is hosted in Australia by a hosting service (see [Who is in scope?](#)).



Who is in scope?

The BSA and AVMA generally apply to **ISPs, hosting service providers, and content service providers**, although these entities are defined differently under both regimes. This generally includes **social media platforms and cloud-based storage sites**, although some exclusions apply under the BSA. The EOSA applies to **social media services**, relevant electronic services (such as **emails, text and online messages, chat services and online gaming**) and designated internet services (such as **websites**).

Although **search engines** may be covered by the BSA, the definition is narrow and would not capture general search engines such as Google.

Search engines are not in scope under the AVMA or EOSA.



What do you have to do to comply?

Under the BSA, providers must:

- > **comply with any takedown notice, service cessation notice or link-deletion notice** issued by the Australian eSafety Commissioner, as well as any directions given on how to remediate the issue (for example, by effectively filtering out the content through software); and
- > **comply with the relevant online provider rules and industry standards**, including the Australian Internet Industry Code of Practice. These rules outline a

minimum standard for the use of content filtering software, including “family friendly content” filtering.

Under the AVMA:

- > ISPs, providers of content, and hosting services providers must **notify the police within a reasonable time of becoming aware** that their service can be used to access abhorrent violent material, or that recording or streaming of such material has occurred or is occurring in Australia. What is a reasonable time will depend on the type and volume of the material and the capabilities available to the provider; and

- > content and hosting services providers must **ensure the expeditious removal of abhorrent violent material** from their service.

Under the EOSA:

- > social media services must have **terms of use which prohibit users from posting cyberbullying material** and must maintain a complaints scheme for users to request removal of cyberbullying material. The services must comply with a **user’s request for removal** under the complaints scheme within 48 hours; and
- > users and providers of a social media service, relevant electronic service or designated internet service must comply with notices from the eSafety Commissioner requiring the **removal of intimate images from the service within 48 hours**. Hosting service providers must also comply with notices to cease hosting the intimate image within 48 hours.



What types of harm are regulated?

The **BSA covers a wide range of harmful content**. It uses the concept of “prohibited content”, meaning content designated by the Classification Board as offending the standards of morality, decency and propriety as generally accepted by reasonable adults, including terrorist material. There is also “potential prohibited content”, which covers material which has not been considered by the Classification Board yet, but is substantially likely to be considered prohibited content.

The **AVMA covers a more limited range of harmful content**. It regulates “abhorrent violent material”, which is limited to offensive audio, visual (including videos and photos), or audio-visual material which records or streams violent conduct, such as terrorism, murder or torture.

The EOSA covers **cyberbullying material** on social media services or relevant electronic services generally and cyberbullying material targeted towards a child. Cyberbullying material targeted towards a child includes any material which would likely cause the child to feel seriously threatened, seriously intimidated, seriously harassed or seriously humiliated. The EOSA also covers intimate images of a person performing sexual, intimate, or private activities shared on a hosting service. An intimate image includes where a person, due to their religious or cultural background, wears a particular attire and the image depicts them without that attire.



Does the regime cover private communications?

Yes. Under the BSA, an individual or organisation can report private communications which contain (potential) prohibited content, and the regulator can investigate these reports and take enforcement action.

Similarly, if an entity regulated under the AVMA discovers that abhorrent violent material is accessible on their platform, they must remove or report this content, even if the communications are private.

Under the EOSA an individual can report cyberbullying material which occurs in private communications. An individual can also report if a person posts or threatens to post an intimate image of the individual through private communications without the individual’s consent.



Who is the regulator, and how do they enforce the regime?

The **Australian eSafety Commissioner** is responsible for enforcing Schedules 5 and 7 of the BSA, the AVMA and the EOSA.

Under the BSA:

- > if potential prohibited content is being hosted in, or provided from, Australia, the eSafety Commissioner will issue the provider with an interim take-down notice or link-deletion notice. The provider must comply as soon as practicable, and by no later than 6pm the next business day.

- > if potential prohibited content is being hosted outside Australia, the eSafety Commissioner must notify law enforcement and ISPs. The ISP must take all reasonable steps to prevent users from accessing the content as soon as practicable, and by no later than 6pm the next business day.

The BSA sets out a **range of criminal, civil and administrative sanctions** for failure to comply, including:

- > if potential prohibited content is being hosted in, or provided from, Australia, criminal and civil offences for contravening an online provider rule (including a failure to comply with a take-down or link-deletion notice, industry standard or direction of the eSafety Commissioner), with **a fine of up to AUD 11,100 per day for individuals and up to AUD 55,500 per day for companies**;
- > if potential prohibited content is being hosted in, or provided from, Australia, criminal and civil offences for contravening a designated content / hosting service provider rule (including a failure to comply with a take-down or link-deletion notice, industry standard or direction of the eSafety Commissioner), with **a fine of up to AUD 22,200 per day**;
- > if potential prohibited content is being hosted outside Australia, criminal and civil offences for breach of a content / hosting service provider rule (including a failure to comply with an access-prevention notice, industry standard or direction of the eSafety Commissioner), with **a fine of up to AUD 11,100 per day for individuals and up to AUD 55,500 per day for companies**;
- > **remedial directions and warnings**; and
- > **a court order** to stop the provider supplying an internet carriage service or providing the designated content / hosting service.



Under the AVMA:

- > if a regulated entity fails to report abhorrent violent material to law enforcement, they can incur **a fine of approximately AUD 177,000**; and
- > for failure to remove abhorrent violent material expeditiously, a regulated entity can face a **fine of the greater of AUD 11.1 million or 10% of annual turnover**, and an **individual may be subject to a fine of AUD 2.22 million and up to three years' imprisonment** where they have provided the content or hosting service.

Under the EOSA:

- > if a social media service is requested to remove cyberbullying material by a user under its complaints scheme and the social media service does not comply with this request within 48 hours, the social media service may be issued a warning or a **fine of up to AUD 111,000**; and
- > if the eSafety Commissioner is satisfied that an intimate image posted without the individual's consent is hosted by a hosting service, the eSafety Commissioner may issue the hosting service a removal notice to remove the image within 48 hours. If the hosting service does not comply with the removal notice, the service may be issued a warning or a **fine of up to AUD 111,000**.



Looking ahead

Online Safety Bill

In February 2021, the Online Safety Bill was introduced into Parliament. This Bill proposes major reforms to Australia's current regulation of online harms by significantly consolidating and modernising the currently separate regimes.

In particular, the Bill:

- > introduces **basic online safety requirements** for social media services, relevant electronic services and designated internet services, as well as mandatory reporting requirements and penalties for failure to comply;
- > introduces a **cyberbullying scheme for adults** which provides the eSafety Commissioner with the power to issue 24-hour removal notices (compared to the current 48-hour removal notices);
- > incorporates the current regime in the EOSA which requires the **removal of intimate images shared without a person's consent**, reducing the take-down time for such material to 24 hours (compared to the current 48-hour removal notices); and
- > **simplifies and updates the online content scheme** in Schedules 5 and 7 of the BSA, and provides the eSafety Commissioner with the power to request or determine an industry code which must be complied with.

Australian Code of Practice on Disinformation

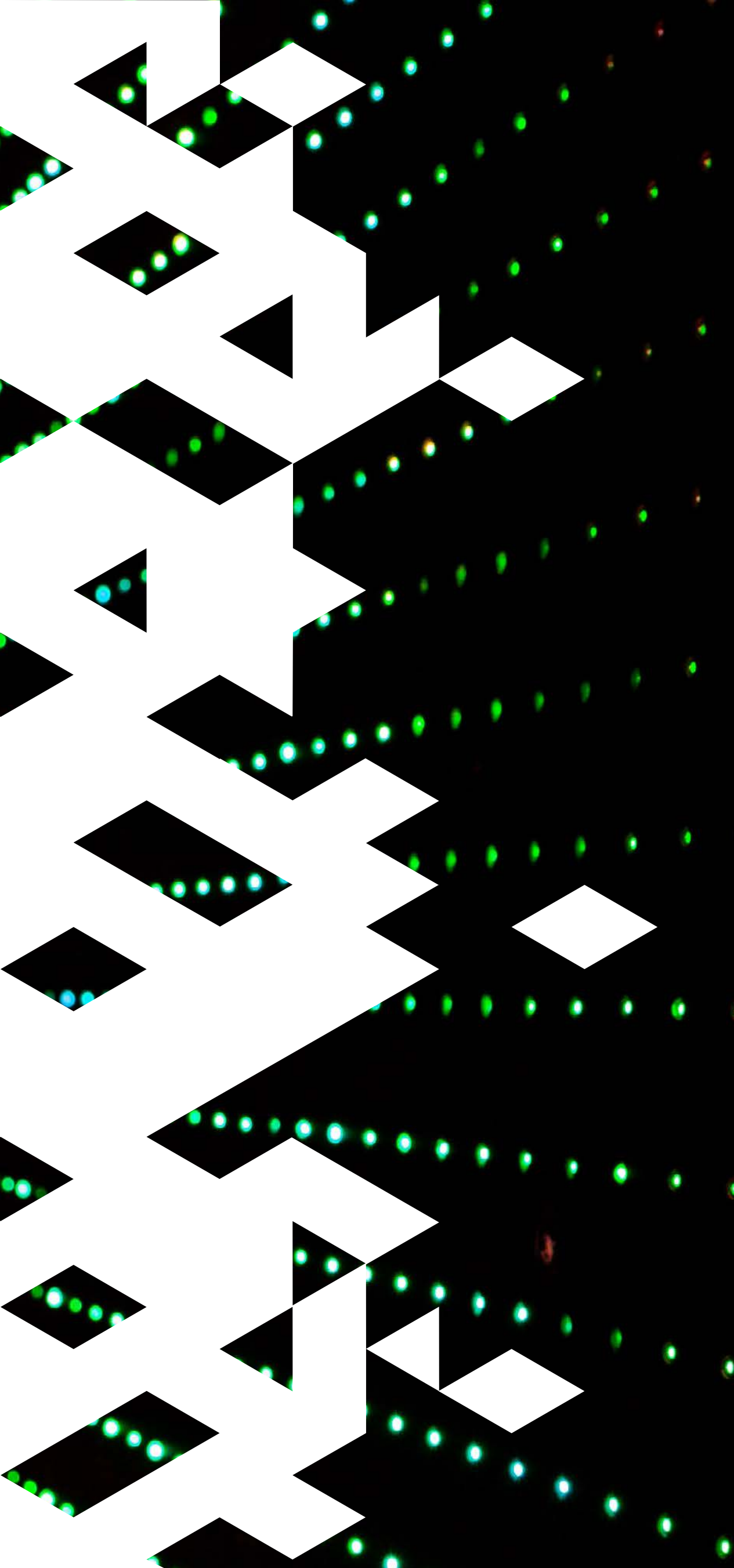
In addition, the Digital Industry Group Inc released the "Australian Code of Practice on Disinformation" on 22 February 2021. Industry participants may voluntarily sign the Code and commit to the Code objectives that are relevant to their business. The Code includes objectives for signatories to:

- > develop and implement scalable measures to reduce the risk of harms resulting from misinformation and disinformation on their platforms;
- > develop and implement a reporting process for users to report misinformation and disinformation on their platforms; and
- > prepare an annual transparency report setting out their progress towards the objectives in the code.

As at **February 2021, Facebook, Google, Microsoft, Redbubble, TikTok and Twitter have signed the Code** and have committed to releasing annual transparency reports.

A sub-committee will meet every six months to monitor how signatories are meeting their commitments.





European Union



What framework is in place to regulate online harms?

e-Commerce Directive

Currently, the regulation of digital services within the EU is governed by the e-Commerce Directive, which harmonised the cross-border provision of online services in the EU when it was adopted in 2000.

Digital Services Act

In recognition of the evolution of the digital sector since 2000, and in response to pressure from various stakeholders, in December 2020 the European Commission proposed a significant legislative reform of the EU digital economy via the Digital Services Act (the “**DSA**”).

Audiovisual Media Services Directive

Some forms of online harm are also regulated by the Audiovisual Media Services (“**AVMS**”) Directive, after a 2018 amendment expanded its scope to cover video sharing platforms (“**VSPs**”). VSPs must protect the general public from several specific types of online harm.

Given the limited scope of the AVMS Directive, it is not considered further in this note. Instead, we have focused our analysis on the DSA proposals, given their relevance for platforms as a whole.

For more information on the implementation of the VSP regime in the UK, please see our articles [The UK’s VSP regime: flight-testing the regulation of social media firms](#) and [Learn to fly: Ofcom releases the flight manual for the VSP regime](#).

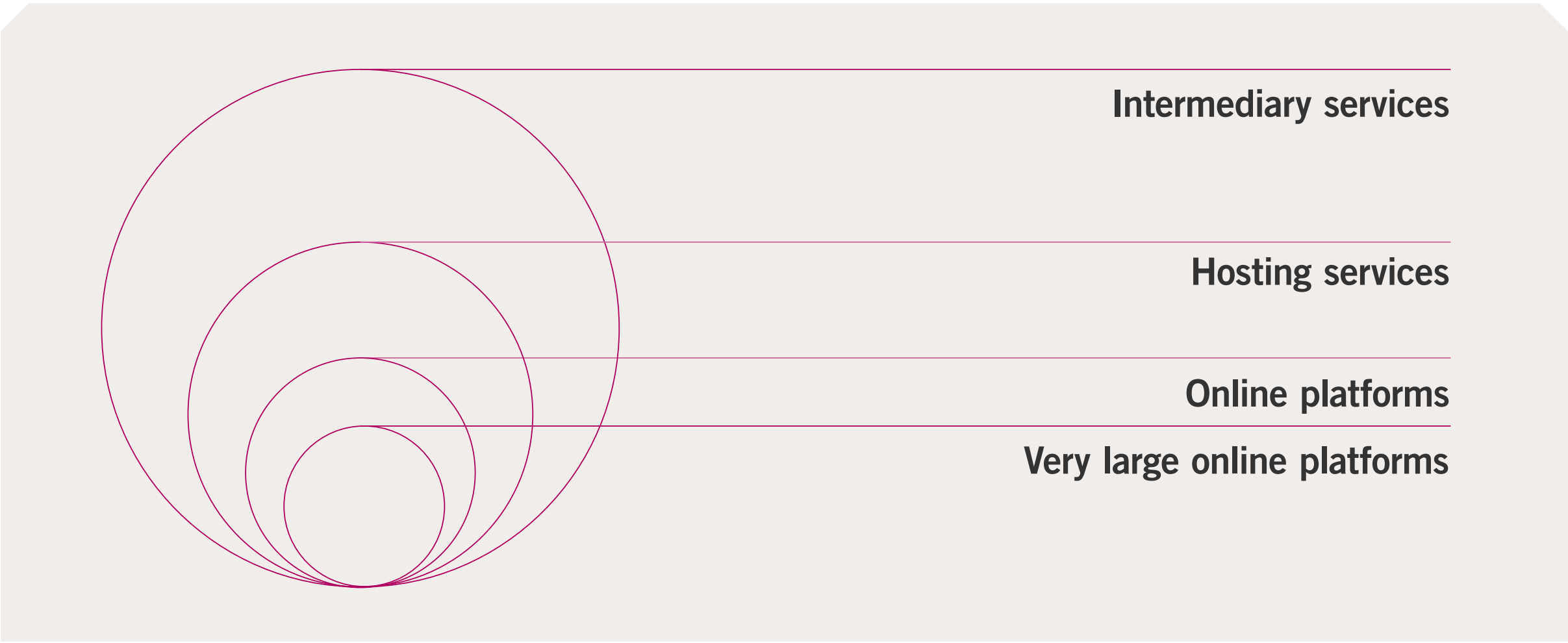


Who is in scope?

The DSA applies to all “online intermediary services”, for example, internet access providers and domain name registrars. Within that definition, there are subcategories of:

- > **hosting services** (for example, cloud and webhosting services), which includes
- > **online platforms** which store and disseminate information to the public (for example, social networks, content-sharing platforms, app stores, online marketplaces, online travel and accommodation platforms), which includes
- > **very large online platforms** (“**VLOPs**”), systemic platforms with at least 45 million monthly active users in the EU.

All online intermediaries offering their services in the EU, whether they are established in the EU or outside, will have to comply with the new rules.





What do you have to do to comply?

The DSA **retains the exemption from liability set out in the e-Commerce Directive** that a hosting platform is only liable for the content on its site if it has actual knowledge that this content is illegal and fails to act expeditiously to remove it or disable access to it.

It builds on this exemption by clarifying that intermediaries choosing to carry out voluntary own-initiative investigations will not be deprived of the benefit of the exemption.

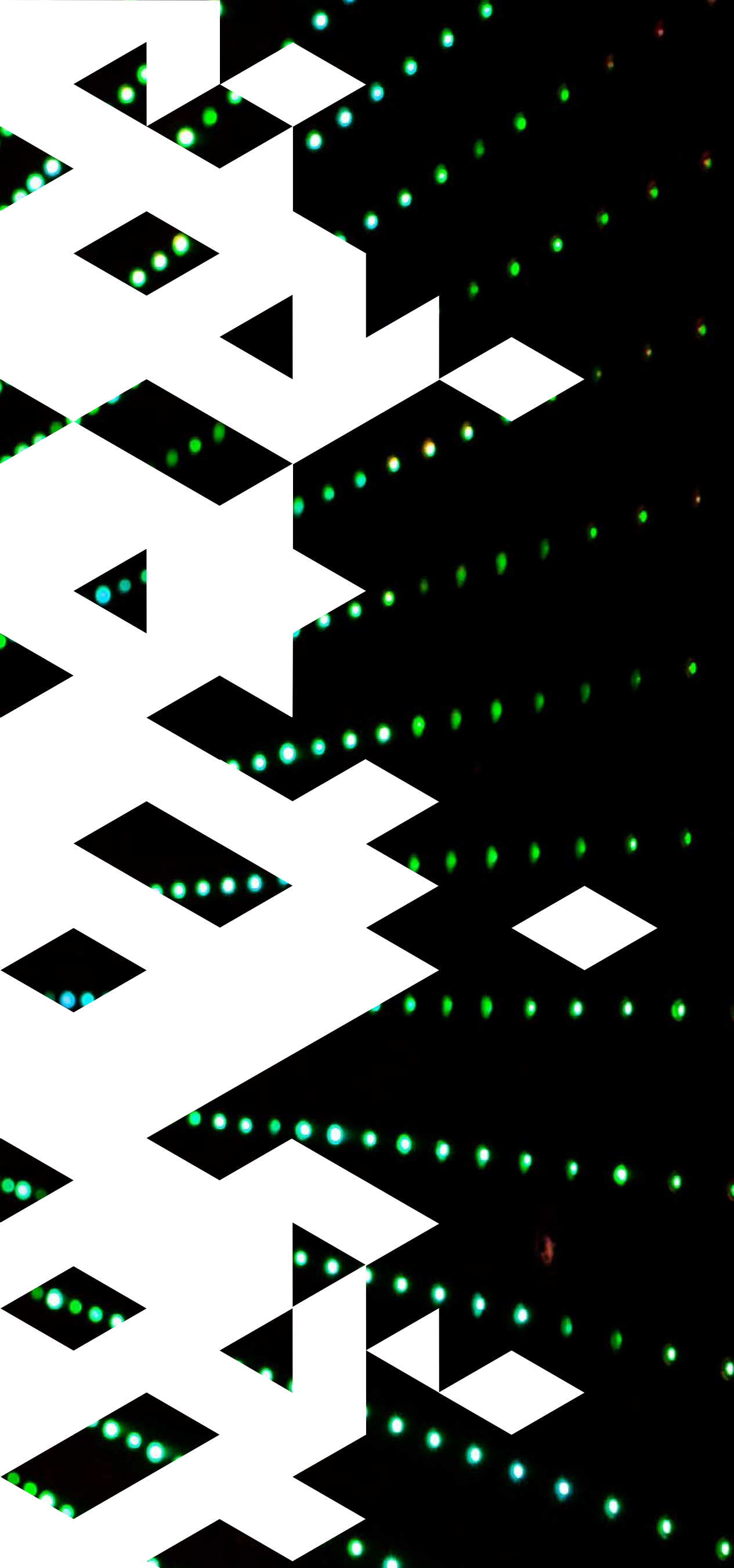
However, the DSA imposes **new obligations on digital service providers** centred around four main principles:

- > **Transparency:** All digital service providers without an establishment in the EU must appoint a legal representative in a Member State where they offer services. They must publish clear, comprehensible and detailed annual reports on content moderation (with additional information required for online platforms and VLOPs). Online platforms must also ensure that traders provide sufficient information to the platform and display trader information to users. Online platforms must provide transparency on advertisements and on the algorithms used to display them (with additional requirements for VLOPs). VLOPs must also publish information on their use of recommender systems.
- > **Empowering users:** All digital service providers must include information on any content restrictions that they impose in their terms and conditions. Providers of hosting services must set up a notice mechanism for users to report illegal content and they must give a statement of reasons when they remove or disable access to specific content. Online platforms must provide content dispute resolution mechanisms enabling users to appeal their decisions.

- > **Risk management:** Online platforms must take measures to protect their systems against misuse, including obligations to remove illegal goods, services or content. Online platforms must inform the relevant authorities if they suspect a serious criminal offence involving a threat to the life or safety of persons. VLOPs must also take steps to manage systemic risks, including annual risk assessments, risk mitigation measures, annual independent audits and appointing compliance officers.
- > **Industry co-operation:** The European Commission will support and promote the development of voluntary industry standards, codes of conduct and crisis protocols on certain aspects of online businesses.

The obligations which apply will depend on the type of digital service provider, with more extensive obligations for online platforms and VLOPs.

Obligation	Intermediary services	Hosting services	Online platforms	VLOPs
Transparency				
Appoint a point of contact/ legal representative	✓	✓	✓	✓
Transparency reports	✓	✓	✓	✓
Transparency in online advertising			✓	✓
Transparency in trader information			✓	✓
Transparency in recommender systems				✓
Empowering users				
Information on content restrictions in Ts and Cs	✓	✓	✓	✓
Notice and action mechanism		✓	✓	✓
Content dispute resolution mechanism			✓	✓
Risk management				
Report criminal offences			✓	✓
Annual risk assessment and audit				✓
Risk mitigation measures				✓
Compliance officer				✓



What types of harms are regulated?

All intermediary services must take specific measures to protect their systems against misuse, including obligations to remove illegal goods, services or online content, as set out in the previous table. However, there are **additional requirements for hosting services, online platforms and VLOPs**.

In particular, VLOPs have specific obligations in relation to certain types of harmful content. They must assess, and take steps to mitigate, systemic risk to users of their service in relation to:

- > the **dissemination of illegal content**;
- > **negative effects for the exercise of fundamental rights**, for example the right to private and family life, freedom of expression etc.; and
- > **intentional manipulation of their services with an actual or foreseeable effect** on public health, minors, civic discourse, electoral process or public security.

Measures that may be needed to address these risks could include adapting content moderation or recommender systems, discontinuing advertising revenue for specific content, and improving the visibility of authoritative information sources.

The concept of “**illegal content**” is defined broadly and refers to information that under the applicable law (EU and/or relevant Member State) is either itself illegal, such as illegal hate speech or terrorist content, or relates to activities that are illegal, such as the sharing of images depicting child sexual abuse.



Does the regime cover private communications?

Potentially. Recital 14 of the proposals states that “*Interpersonal communication services, as defined in Directive (EU) 2018/1972 of the European Parliament and of the Council, such as emails or private messaging services, fall outside the scope of this Regulation*”. However, to the extent that these private communication services also double as an intermediary service (or another category of provider), they may nevertheless also become subject to the DSA – the proposals are unclear on this point. This should hopefully be clarified as part of the EU’s co-decision process.



Who is the regulator, and how do they enforce the regime?

Digital Services Coordinator

Each Member State will be required to appoint a Digital Services Coordinator (“**DSC**”) to enforce the DSA.

DSCs will have various powers of investigation, including to carry out on-site inspections, interview staff members and require the production of documents and information.

If a DSC finds that a digital service provider has breached its obligations, it will have the power to:

- > **order the cessation of infringements**;
- > **impose interim measures**; and
- > **impose fines of up to 6% global annual turnover**, or periodic penalty payments of up to 5% of average global daily turnover.

European Commission

In cases concerning VLOPs, the issue can be escalated to the Commission. The Commission will have the same investigatory and enforcement powers as the DSCs.



Looking ahead

The DSA will now go through the EU’s co-decision process through which both the European Parliament and Council must agree on the final text of the legislation.

We expect the DSA to come into force in 2023 at the earliest.



France



What framework is in place to regulate online harms?

Avia Law

Law n°2020-766 of 24 June 2020 on hate speech on the Internet, also called “Loi Avia” (the “**Avia Law**”, named after the law’s main sponsor) came into force on 26 June 2020. Its provisions modify Law n°2004-575 of 21 June 2004 on Confidence in the Digital Economy (“**LCEN**”), which largely mirrors the provisions of the eCommerce Directive.

The Avia Law was intended to be much broader than its current scope. It originally included provisions similar to those found in the NetzDG in Germany. However, many of those provisions were quashed by the French Constitutional Court on 18 June 2020. The court held that a proposed requirement that platforms remove “manifestly” illegal content within 24 hours was incompatible with the right to freedom of expression, given the risk that platforms would “over-block” to avoid enforcement action. This summary therefore focuses on the narrower final scope of the Avia Law.

Digital Services Act

This regulatory framework will also be impacted by the introduction of the DSA (see our section on [the EU](#)). The DSA will include some of the more ambitious measures that were quashed by the French constitutional court, as well as measures going beyond those contemplated even by the initially broad scope of the Avia Law.



Who is in scope?

The regime applies to:

- > **individuals or legal entities whose business is providing online communication services to the public;** and
- > **individuals or legal entities that provide a platform to store text, images, messages, etc. which are made available to the public** (e.g. social media platforms).



What do you have to do to comply?

The LCEN requires that those in scope:

- > set up an **easily accessible system** to allow users to report hate speech;
- > publish **details of the resources** they devote to tackling hate speech on their platforms;
- > **remove child sexual abuse or terrorist content within 24 hours of being notified** of the material by the general directorate of the national police; and
- > **promptly inform the competent public authorities of harmful content reported to them.** The law does not define “promptly”, but case law suggests that a delay of five days is too long. The platform must also provide any data they hold which would help to identify the user who posted the content.



What types of harm are regulated?

The LCEN provides an exhaustive list of “**hate speech content**”, i.e. anything that breaches the French Criminal Code or the French Law on the Freedom of the Press of 29 July 1881.

This includes:

- > sexual harassment;
- > provoking hatred or violence against a person based on their gender, sexual orientation, disability, race, or religion; and
- > directly provoking or condoning terrorist acts.



Does the regime cover private communications?

No.



Who is the regulator, and how do they enforce the regime?

Courts and French general directorate of the national police

The proposed Avia Law appointed the Superior Council for Audio-visual (the “**CSA**”) as the regulator, but this provision was quashed by the French Constitutional Court. The courts and the French general directorate of the national police will instead be responsible for enforcing the provisions.

The Avia Law also created a research body to monitor and analyse the development of online hate speech.

The courts have broad powers to enforce the regime, including orders to block access to certain websites.

The French general directorate of the national police has the power to demand the removal of child sexual abuse or terrorist content.

Criminal law sanctions

The Avia Law introduced higher fines for criminal non-compliance. Criminal law sanctions can include:

- > **for individuals, including directors and senior employees of providers, fines of up to EUR 250,000 and one year’s imprisonment.** In practice it is rare for individuals to be held criminally liable; and
- > **for companies, fines of up to EUR 1.25 million and a prohibition preventing them carrying out their activity for five years.**

Damages and liability

Damages for civil liability are based on French tort law which requires that any damage to the claimant is fully repaired.

The LCEN states that providers will only be criminally and civilly liable if they have actual knowledge of the harmful content and do not act promptly to remove or block the content.

The law also states that websites which carry terrorist or child sexual abuse content can be blocked and removed from search engine results.



Looking ahead

A few days after the French Constitutional Court’s decision in relation to the Avia Law, the French government asked the European Commission to adopt a new law to force platforms to remove illegal content by appointing an independent regulator with the power to issue binding recommendations and impose sanctions.

The French government welcomed the recent publication of the DSA, which it considers to be in line with its request.



Germany



What framework is in place to regulate online harms?

NetzDG

The Act to Improve Enforcement of the Law in Social Networks (“**NetzDG**”) came into force in 2017, was last amended in April 2021 by the Act to Combat Right-Wing Extremism and Hate Crime, and is aimed at tackling hate crime and other harmful content on social networks.

Digital Services Act

It is not yet clear how the DSA proposals will affect the current regulatory framework in Germany (see our section on [the EU](#)). The DSA proposals which exceed the current scope of the NetzDG will replace the relevant provisions in Germany.



Who is in scope?

The NetzDG applies to the “**providers of social networks**”, but **not to their users**.

“Providers of social networks” are service providers who, for profit, operate internet platforms designed to allow users to share content, regardless of where they are established. “Content” includes own and third-party content, such as images, video and text.

Some **key exemptions** from this definition include providers whose platforms:

- > have fewer than two million registered users in Germany;
- > offer journalistic content which the service provider is responsible for; or
- > are designed to enable individual communication (e.g. email or messenger services) or the dissemination of specific content. The term “dissemination of specific content” is imprecise, but an explanatory memorandum to the NetzDG states that business networks (such as LinkedIn), video games, marketplaces and communication platforms for discussion of specific issues are out of scope. However, a far-right platform would not be able to benefit from this exemption if it were to limit discussion to “refugees”, for example.



What do you have to do to comply?

Providers must implement effective control mechanisms to filter, block or take down unlawful content on their platforms. This means:

- > **having an effective and transparent system for managing user reports;**
- > offering users an easily accessible **way to report any unlawful content;**
- > **reviewing any reported content expeditiously.** “Manifestly unlawful” content must be blocked or removed within 24 hours. Any other unlawful content must be blocked or removed within seven days;

- > **report harmful content to the Federal Criminal Police Office** (as of February 2022) if the content meets the definition of certain criminal offences; and
- > a range of other **reporting and transparency obligations**. For example, providers who receive more than 100 reports about unlawful content per calendar year must publish bi-annual reports on their reports handling process.



What types of harm are regulated?

The NetzDG regulates “**unlawful**” content, i.e. any content which is punishable under specific criminal offences in the German Criminal Code (“**GCC**”), and which cannot be “justified” (i.e. content which cannot be justified under the “safeguarding of legitimate interests” exception in the GCC).

The NetzDG contains an exhaustive list of the relevant criminal offences, including:

- > **use of symbols of ‘unconstitutional’ organisations** (section 86a GCC), such as far-right and/or terrorist organisations;
- > **public incitement to commit offences** (section 111 GCC); and
- > **forming criminal or terrorist organisations** (sections 129 – 129b GCC).

The NetzDG is primarily focused on the regulation of hate crimes and fake news. Though spreading untrue factual claims is not a specific offence under the GCC, doing so could fall under one of the other offences listed and therefore constitute unlawful content.



Does the regime cover private communications?

No. The NetzDG only applies to content made available to the public on a social network. It does not cover platforms that enable private communication between two or more users. This means that messaging services and data transfer services may fall outside the scope of the NetzDG.



Who is the regulator, and how do they enforce the regime?

Federal Office of Justice

The Federal Office of Justice is the administrative body with the power to enforce the NetzDG. It has oversight of the user reports management and reporting obligations of providers.

Fines

It can impose regulatory fines for non-compliance with certain obligations imposed by the NetzDG, for example if providers do not offer an easily accessible mechanism to report unlawful content.

The Federal Office of Justice cannot decide on the unlawfulness of content itself, because of the division of competences in the German Constitution. If it intends to impose fines because unlawful content has not been removed or blocked, a court must first determine the content’s unlawfulness.

Following the court’s decision, the fine can be imposed by the Federal Office of Justice.

The Federal Office of Justice has discretion to decide on the level of fines, but must follow specific fining guidelines approved by the Federal Ministry of Justice and Consumer Protection for the NetzDG. The fines may amount to:

- > **max. EUR 5 million against the provider’s representatives** (for example, the managing director, or the owner), where they are responsible for the non-compliance; and
- > **max. EUR 50 million against the legal person or association of persons operating the platform.** Because the fine should exceed the economic advantage of the platform, the fine may also exceed the EUR 50 million in specific cases where the platform’s economic advantage is higher than EUR 50 million.

Liability and appeals

The NetzDG does not provide for any criminal liability for non-compliance.

The individual or legal persons may appeal against these fines to the Federal Office of Justice, subject to a two-week time limit. If the Federal Office of Justice upholds the fine after an appeal, the Regional Court of Bonn can then hear the appeal.



Looking ahead

Amendments to NetzDG

In April 2021, the NetzDG was amended by the Act to Combat Right-Wing Extremism and Hate Crime, **simplifying the prosecution of right-wing extremism and hate crime** offences. The amendment creates, as of February 2022, an obligation for platforms to report certain types of unlawful content, such as online threats, and other information such as the IP address of the respective user to the Federal Criminal Police Office.

The NetzDG will be further amended by an act **strengthening the rights of users of social networks** by making reporting channels more user-friendly and creating more transparency by expanding the scope of the biannual reports and creating a right for both the users and the individuals/groups who reported the content to appeal against decisions of the platform not to block or remove reported content. The act has been approved but is not yet in force because of ongoing discussions on its constitutionality.

Digital Services Act

The DSA proposals cover certain aspects of both amendments, such as the obligation to publish regular transparency reports and report serious criminal offences to the authorities. The NetzDG is also stricter in places than the DSA proposals, such as the time limits for the removal of unlawful content.

Judging by the initial political reaction, a careful balance between the harmonising European obligations for platforms stipulated by the DSA and the additional national enforcement stipulated by the NetzDG will be required.

Ireland



What framework is in place to regulate online harms?

Online Safety and Media Regulation Bill

Ireland does not currently have specific legislation governing online harms. However, in January 2020 a draft General Scheme of the Online Safety and Media Regulation Bill was published. The legislation will create a new framework to regulate harmful online content.

Digital Services Act

This framework will also be impacted by the introduction of the DSA (see our section on [the EU](#)).



Who is in scope?

The proposals apply to:

- > **“relevant online services”** (any information society service established in Ireland that allows a user to disseminate or access user-generated content); and
- > **“designated online services”** (any relevant online service that has been designated as such by the Media Commission).

This means that a wide range of different organisations may come within the scope of the new law, including VSPs, social media providers, e-commerce services, online search engines and ISPs.

Different regulations will apply depending on whether the company is a “relevant” or “designated” online service.



What do you have to do to comply?

The proposals focus on the systems and processes that providers have in place, rather than on regulating individual pieces of content.

Providers will have to comply with **Online Safety Codes** prepared by the Media Commission. These codes may require providers to take actions like:

- > **minimising the availability of harmful online content;**
- > implementing **measures in relation to commercial communications available on their services**. The definition of commercial communications is adapted from the AVMS Directive: *“information conveyed by a media service or relevant online service which is designed to promote, directly or indirectly, the goods, services or image of the natural or legal entity pursuing an economic activity”*;
- > **putting in place mechanisms to handle user complaints and issues;**

- > carrying out **risk and impact assessments** in relation to the availability of harmful online content on their services; and
- > **reporting obligations** regarding compliance with the Online Safety Codes.



What types of harm are regulated?

The proposals set out four categories of harmful online content:

- > **material which it is a criminal offence to disseminate under Irish or EU law** (for example, child sexual abuse material or terrorist content);
- > **cyber-bullying material;**
- > **material encouraging or promoting eating disorders;** and
- > **material encouraging or promoting self-harm or suicide.**

The Media Commission will be able to include or exclude other categories of harm.

The proposals also suggest that the Media Commission will have the power to issue online safety guidance materials in relation to “age inappropriate content” (i.e. material that may not necessarily be harmful, but which is unsuitable for minors, such as gratuitous violence or pornographic material).



Does the regime cover private communications?

Yes, the proposed regime will include private communication services as “relevant online services”. They may also be selected by the Media Commission as a “designated online service”.

However, the proposals provide that private communications providers will not be required to follow an Online Safety Code in relation to material that it is legal to disseminate (i.e. legal but harmful content).



Who is the regulator, and how do they enforce the regime?

Media Commission

The outline proposes the establishment of a multi-person Media Commission, including an Online Safety Commissioner. This new body will replace the Broadcasting Authority of Ireland and will also be responsible for the regulation of on-demand services, including radio, television, and video-on-demand services.

The Media Commission will have the power to impose levies on regulated firms to cover the cost of regulation.

Powers

The Media Commission will have the power to:

- > designate online services for regulation;
- > prepare, monitor and conduct investigations into compliance with the Online Safety Codes;

- > audit any complaints or issues handling processes;
- > operate a “super complaints” system, where nominated bodies can bring systemic issues to the Media Commission’s attention; and
- > direct online services to make changes to their systems, processes, policies and design.

Enforcement

The Media Commission may have a broad range of enforcement powers, including:

- > **issuing information requests, compliance notices, and warning notices mandating compliance** (which can be published). In Ireland, failure to comply with an information request or warning notice is a criminal offence, **punishable by a fine of up to EUR 5,000 and/or 12 months’ imprisonment** (on summary conviction). The criminal sanction for conviction on indictment has not yet been published;
- > **pursuing civil sanctions, including administrative fines of up to EUR 20 million or 10% of relevant turnover (whichever is higher) for the preceding financial year**, orders compelling compliance with warning notices, or requiring ISPs to block access to the offending online service in Ireland. These sanctions will require court approval; and
- > **prosecuting summary criminal offences.**

Criminal sanctions

The outline also proposes extending criminal liability to senior management for specified offences, committed by online services, where it is proven that the offence was committed with the consent or connivance of senior management, or where they have been acting with wilful neglect.

At the moment, the only specified offence for which senior management may be in scope for criminal liability is non-compliance by a regulated entity with a warning notice.

Any proceedings, including summary proceedings, for this offence must be instituted by, or with the consent of, the Director of Public Prosecutions.



Looking ahead

The draft legislation is due to undergo pre-legislative scrutiny by the Irish government. It is expected to be enacted by the end of 2021 or 2022 at the earliest.

The Irish government have welcomed the publication of the DSA proposals. It remains to be seen what impact the DSA will have on the Irish government’s current proposals for regulation of online harms.

However, the DSA and Irish proposals take a similar systemic approach to online content regulation. In particular, both proposals focus on identifying risks to users of online services, developing targeted measures to minimise those risks, and reviewing the effectiveness of those measures over time.

Singapore



What framework is in place to regulate online harms?

Broadcasting (Class Licence) Notification

The Broadcasting (Class Licence) Notification regulates prohibited harmful content in Singapore. Those within scope are required to abide by the conditions in the Internet Class Licence and to ensure that content on their platforms complies with the Internet Code of Practice, introduced in October 2016.

Protection from Online Falsehoods and Manipulation Act

In addition, the Protection from Online Falsehoods and Manipulation Act (“**POFMA**”) came into effect in 2 October 2019. It is also known as the “fake news law”.



Who is in scope?

The Internet Code of Practice applies to **ISPs and internet content providers (“ICPs”)**. ICPs are defined broadly to mean any individual in Singapore who provides any programme online for business, political or religious purposes, including social media platforms.

POFMA applies more broadly to **any users of online services in Singapore**.

The prohibition against the communication of falsehoods does not apply to internet intermediaries, telecoms services, internet access services and computing resource services.

However, the government may still issue directions under POFMA to internet intermediaries.



What do you have to do to comply?

Under the Internet Code of Practice, ICPs and ISPs must use **best efforts to ensure that “prohibited material” is not broadcast via the internet to users in Singapore**. Where providers have no editorial control, the Internet Code of Practice requires them to block access to any “prohibited material”, if directed to do so by the Infocomm Media Development Authority of Singapore (“**IMDA**”).

For providers, POFMA is focused on reactive measures where a falsehood is published, rather than systems and controls to prevent falsehoods being published in the first place.

However, prescribed internet intermediaries are subject to various Codes of Practice under the POFMA, which outline **measures that prescribed internet intermediaries are required to implement to prevent the misuse of online accounts**.



What types of harm are regulated?

The Internet Code of Practice covers **material that is deemed to be objectionable on the grounds of public interest, morality, public order, security, national harmony, or is otherwise prohibited under Singapore law (“prohibited material”)**.

Factors to be taken into account include whether the material depicts extreme violence or incites ethnic or religious intolerance. The Internet Code of Practice **defines “objectionable” content broadly** and it does not specifically limit such content to illegal content.

POFMA regulates **“falsehoods” which are likely to be prejudicial to the security, health, or safety of Singapore, to damage Singapore’s relations with other countries, or to influence the outcome of an election**, among other things.



Does the regime cover private communications?

Yes. The Internet Code of Practice refers to private discussion forums hosted on ICPs’ services, like chat groups. POFMA covers all communications of false statements of fact and does not distinguish between private and public communications.



Who is the regulator, and how do they enforce the regime?

IMDA

The regulator for the Internet Code of Practice is the IMDA, a statutory board of the Singapore government within the Ministry of Communications. It may direct an ICP to deny access to “prohibited material”, and may impose fines at its discretion (including quantum) where an ICP fails to comply with the Internet Code of Practice. ICPs may appeal the IMDA’s decision to the Minister, whose decision will be final.

Enforcement

Government ministers are responsible for the enforcement of POFMA. The legislation also provides for the appointment of an alternative authority during election periods and other specified periods. The POFMA Office within the IMDA is responsible for the administration of POFMA.

To enforce POFMA, government ministers can issue **a wide range of directions, requiring providers to put up a notice that a communication is false, remove the communication, or, in the case of ISPs, to disable end-user access to the relevant statements.** Ministers have broad enforcement powers to conduct investigations, and POFMA provides for a range of different fines and offences for non-compliance with ministerial directions. **Individuals can be fined up to SGD 20,000 and/or imprisoned for up to 12 months, and companies can be fined up to SGD 1 million.** Ministers can also make an access blocking order.

POFMA provides that an individual will not incur civil or criminal liability for an act or omission where it was done with reasonable care and in good faith for the purpose of complying with a direction.



Looking ahead

Singapore continues to develop its regulation in this space.

The POFMA Office has issued various Codes of Practice to ensure internet intermediaries have adequate systems and processes in place to prevent the misuse of online accounts, such as:

- > the Code of Practice for Preventing and Countering Abuse of Online Accounts;
- > the Code of Practice for Transparency of Online Political Advertisements; and
- > the Code of Practice for Giving Prominence to Credible Online Sources of Information (which came into force on 2 October 2019).





United Kingdom



What framework is in place to regulate online harms?

Online Safety Bill

Aside from the handful of companies within the scope of the VSP regime (see our section on [the EU](#)), the UK does not currently have specific legislation governing online harms.

However, in December 2020 the UK government published its final proposals for a new regulatory regime aimed at protecting UK users against harmful online content which it hopes will make the UK “*the safest place in the world to go online*”.

Draft legislation, the Online Safety Bill, is expected to be published in 2021.



Who is in scope?

With only a few exceptions, the new regime will apply to all companies that:

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

In-scope services will include **social media platforms, consumer cloud storage sites, video sharing platforms, online forums, video games which enable interaction with users online, and online marketplaces**.

Search engines will also be in the scope of the regime, despite not hosting user-generated content directly and/or enabling user interaction.



What do you have to do to comply?

Statutory duty of care

In-scope platforms will have to comply with a statutory duty of care to “*take action to prevent user-generated content or activity on their services causing significant physical or psychological harm to individuals*”.

In practice, platforms will fulfil this duty of care by:

- > **conducting risk assessments** to understand the risk of harm to individuals using their services; and
- > **putting in place appropriate systems and processes** to improve user safety by reducing the risk of harms they have identified occurring.

All in-scope platforms will be under a duty to protect:

- > **all users against illegal content and activity** – platforms must assess the nature and level of risk of illegal harms on their services, take steps to prevent the use of their services for criminal activity, and minimise the risk of illegal content appearing on their services; and
- > **children from harmful content** – platforms must assess whether children are likely to access their services. If the platform considers that children are likely to access their services, they will be required to conduct a child safety risk assessment and identify and implement proportionate mitigations to protect children.

All in-scope platforms must also have **effective user reporting and redress mechanisms in place**.



Additional duties for high-risk and high-reach

Services which are considered to be “*high-risk and high-reach*” (so-called “Category 1 companies”) will be under additional duties in relation to content and activity that is **legal but harmful to adults**. They must complete regular risk assessments to identify legal but harmful materials on their services, set clear and accessible terms and conditions which state how they will handle such material, and enforce those terms and conditions consistently and transparently.



What types of harm will be regulated?

The UK government has defined harmful content as content that “*gives rise to a reasonably foreseeable risk of a significant adverse physical or psychological impact on individuals*”.

The legislation will not set out an exhaustive list of harms, but secondary legislation will set out a **limited number of priority categories of harm**, including:

- > **priority categories of criminal offences**, such as child sexual exploitation and abuse, terrorism, hate crime and the sale of illegal drugs and weapons;
- > **priority categories of harmful content and activity affecting children**, such as pornography and violent content; and
- > **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. Some types of disinformation and misinformation are also likely to be included.



Will the regime cover private communications?

Yes. As noted above, the regime will apply where platforms facilitate private user interactions.

The UK government has not yet provided full detail on exactly what it expects in relation to private communication channels.



Who is the regulator, and how will they enforce the regime?

Ofcom

The UK government has confirmed that Ofcom, the existing communications industry regulator, will be the regulator for the online harms regime.

The legislation will impose various obligations on Ofcom, for example it will be required to:

- > **publish codes of practice** which outline the systems and processes platforms need to adopt to fulfil the duty of care;
- > **establish a super-complaints function** to address systemic issues affecting a large number of individuals; and
- > **establish appropriate mechanisms for user advocacy.**

If Ofcom finds that a platform has failed to comply with its regulatory obligations under the new regime, it will have a broad range of enforcement options, including the power to:

- > **issue a fine of up to the greater of GBP 18 million or 10% of the platform’s annual turnover;**
- > **require third parties to withdraw access to key services** that make it less commercially viable for the platform to operate within the UK; and
- > require key internet infrastructure service providers to take steps to **block a platform’s services** from being accessible in the UK, for example ISP blocking.

Potential criminal sanctions for senior managers

The UK government has reserved its right to introduce criminal sanctions for senior managers. However, the scope of any such liability will be limited to situations where senior managers fail to respond “*fully, accurately and in a timely manner*” to information requests from Ofcom, rather than liability for a breach of the duty of care itself.



Looking ahead

The UK government will publish the text of the Online Safety Bill later this year and we expect that it will enter into force in 2022/2023.



United States



What framework is in place to regulate online harms?

Section 230

In the U.S., there is no legislation that requires platforms to take measures in respect of harmful content online. Indeed, section 230 of the Communications Decency Act of 1996 (the “**CDA**”) provides that “*no provider or user of an interactive computer service shall be treated as the publisher or speaker of any of the information provided by another information content provider*”.

In effect, this provides online platforms and ISPs with broad immunity from liability for user-generated content on their platforms.

There is no equivalent of the “notice and takedown” requirement found in other online harms regimes. In general, online platforms are immune from liability even if they are made aware of illegal content on their site.

Section 230 also protects platforms where they voluntarily take steps in “good faith” to moderate user-generated content, by ensuring they will not be held liable for their moderation decisions. This is intended to encourage platforms to engage in content moderation without fear of being held liable for these moderation decisions.

Freedom of speech

The U.S.’s approach to online harms regulation is influenced by its focus on freedom of speech, in particular as set out in the First Amendment. A number of commentators have pointed out that, even without section 230, the First Amendment would protect “legal but harmful” online content in the U.S., such as disinformation or hate speech.

Exceptions

There are a number of narrow exceptions to the protections of section 230, including:

- > where the platform itself is deemed to be a “content provider”, and therefore does not fall within the scope of the immunity in section 230;
- > violations of federal criminal law;
- > intellectual property claims; and
- > certain sex trafficking and prostitution offences. In 2018, Stop Enabling Sex Traffickers Act (“**SESTA**”) and Allow States and Victims to Fight Online Sex Trafficking Act (“**FOSTA**”) were passed, introducing a carveout to section 230 for civil and criminal charges of sex trafficking, and conduct that “*promotes or facilitates prostitution*”.

The potential for reform

Section 230 is now 25 years old, and the internet has changed beyond recognition in that time. In the past two years, section 230 has gone from being a provision little known outside of technology legal circles to a topic regularly debated by politicians, the media and the public.

The key question is what reforms, if any, should be made to extend the liability of platforms for user-generated content. In the **Looking Ahead** section, we examine some of the proposals for reform.



Who is in scope?

The protection of section 230 applies to providers and users of an “interactive computer service”, defined as any information service, system, or access software provider that provides or enables computer access by multiple users to a computer server.

In practice, this will apply to **most online platforms, including social media platforms, ISPs, and internet search engines**.



What do you have to do to comply?

As noted above, U.S. law does not impose obligations on online platforms to prevent harmful content appearing on their services.

However, to benefit from the exemption from liability, platforms must not cross the line into becoming a “content provider” or fall within any of the other exemptions outlined in **What framework is in place to regulate online harms?**



What types of harm are regulated?

Section 230 does not regulate specific types of harm.



Does the regime cover private communications?

No.



Who is the regulator, and how do they enforce the regime?

As the U.S. does not have a regulatory regime for online content, there is accordingly no designated regulator either. Debates about whether online platforms are liable for online content tend to be played out in civil litigation pursued by affected individuals instead.

While section 230 provides online platforms with immunity from liability for information provided by another content provider, affected individuals may successfully hold an online platform liable if it acts as something other than “the publisher or speaker of any information provided by another information content provider”.



Looking ahead

Executive Order on Preventing Online Censorship

In May 2020, former President Trump signed an Executive Order on Preventing Online Censorship, which purports to clarify the scope of immunity under section 230. It says “Section 230 was not intended to allow a handful of companies to grow into titans controlling vital avenues for our national discourse under the guise of promoting open forums for debate”.

This Executive Order has been challenged in district courts in the U.S. but these challenges have proved unsuccessful and the order has not yet been revoked by the Biden administration.

Proposed reform of section 230

On 23 September 2020, the U.S. Department of Justice sent draft legislation to Congress to reform section 230. These proposals would limit the immunity provided for by section 230 to content moderation decisions made in good faith and based on an objectively reasonable belief that material promotes terrorism, violent extremism, self-harm or is unlawful. Immunity would not be available where a provider had actual notice of a material or activity on their platform which violates federal law and failed to remove or restrict access to the material, report it to law enforcement where required, or to preserve related evidence.

This draft legislation was not introduced to Congress as a bill, but there is continued bipartisan support for the passing of similar legislation reforming section 230.

SAFE TECH Act

The most recent proposal to reform section 230 by the current U.S. Congress is the SAFE TECH Act, which would introduce a number of changes to section 230, such as removing immunity for any speech a provider had been paid to host, such as advertising or market listing.

It would also add carveouts to section 230 for civil rights laws, stalking and harassment laws, international human rights law, and wrongful death actions.

Biden administration

While President Biden has previously stated that section 230 should be “revoked”, it is likely that his administration will leave the subject to Congress. However, it is still unclear at this stage how the Biden administration plans to approach any reform to online harms regulation.

Focus of U.S. debate

Interestingly, debate in the U.S. has continued to focus on when platforms should be held liable for individual pieces of content rather than considering whether platforms should be required to have certain systems and processes in place.

As systems-focused regimes come into effect across Europe, it will be interesting to see if the U.S. begins to see this as an alternative approach to consider.

Contacts

Australia



Gavin Smith
Partner, Allens
+61 2 9230 4891
gavin.smith@allens.com.au



Angela Kelly
Lawyer, Allens
+61 2 9230 5834
angela.kelly@allens.com.au



Yanery Ventura Rodriguez
Lawyer, Allens
+61 7 3334 3020
yanery.venturarodriguez@allens.com.au

EU Law



Guillaume Couneson
Partner, Belgium
+32 2 501 93 05
guillaume.couneson@linklaters.com



Ceyhun Pehlivan
Managing Associate, Spain
+34 91 399 6182
ceyhun.pehlivan@linklaters.com

France



Sonia Cissé
Counsel, France
+33 1 56 43 57 29
sonia.cisse@linklaters.com

Germany



Julia Schönbohm
Partner, Germany
+49 69 710 03 138
julia.schoenbohm@linklaters.com



Michael Leicht
Partner, Germany
+49 69 710 03 463
michael.leicht@linklaters.com



Lisa Bauer
Associate, Germany
+49 69 710 03 374
lisa.bauer@linklaters.com

Ireland



John Cahir
Partner, A&L Goodbody
+353 1 649 2943
jcahir@algoodbody.com



Davinia Brennan
Associate, A&L Goodbody
+353 1 649 2114
dbrennan@algoodbody.com

Singapore



Adrian Fisher
Partner, Singapore
+65 6692 5856
adrian.fisher@linklaters.com



Jakub Brecka
Associate, Singapore
+65 6692 5897
jakub.brecka@linklaters.com

United States



Adam Lurie
Partner, United States
+1 202 654 9227
adam.lurie@linklaters.com



Charlene Warner
Associate, United States
+1 212 903 9183
charlene.warner@linklaters.com

United Kingdom



Harriet Ellis
Partner, United Kingdom
+44 20 7456 5515
harriet.ellis@linklaters.com



Richard Cumbley
Partner, United Kingdom
+44 207 456 4681
richard.cumbley@linklaters.com



Greg Palmer
Counsel, United Kingdom
+44 20 7456 2925
greg.palmer@linklaters.com



Olivia Grimshaw
Associate, United Kingdom
+44 20 7456 2949
olivia.grimshaw@linklaters.com



Jennifer Calver
Tech Sector Senior PSL, Global
+44 20 7456 2417
jennifer.calver@linklaters.com



Ben Packer
Partner, United Kingdom
+44 20 7456 2774
ben.packer@linklaters.com



Georgina Kon
Partner, United Kingdom
+44 20 7456 5532
georgina.kon@linklaters.com



Jemma Purslow
Managing Associate, United Kingdom
+44 20 7456 4845
jemma.purslow@linklaters.com



Louise Lau
Associate, United Kingdom
+44 20 7456 5375
louise.lau@linklaters.com



Clare Murray
Tech Strategy Consultant, Global
+44 20 7456 2126
clare.murray@linklaters.com



Julian Cunningham-Day
Partner, United Kingdom
+44 20 7456 4048
julian.cunningham-day@linklaters.com



Peter Church
Counsel PSL, United Kingdom
+44 20 7456 5495
peter.church@linklaters.com



Rebecca Dickie
Associate, United Kingdom
+44 20 7456 3324
rebecca.dickie@linklaters.com



Rose Lynch
Associate, United Kingdom
+44 20 7456 2192
rose.lynch@linklaters.com

This publication is intended merely to highlight issues and not to be comprehensive, nor to provide legal advice. Should you have any questions on issues reported here or on other areas of law, please contact one of your regular contacts, or contact the editors.

© Linklaters LLP. All Rights reserved 2021

Linklaters LLP is a limited liability partnership registered in England and Wales with registered number OC326345. It is a law firm authorised and regulated by the Solicitors Regulation Authority. The term partner in relation to Linklaters LLP is used to refer to a member of Linklaters LLP or an employee or consultant of Linklaters LLP or any of its affiliated firms or entities with equivalent standing and qualifications. A list of the names of the members of Linklaters LLP and of the non-members who are designated as partners and their professional qualifications is open to inspection at its registered office, One Silk Street, London EC2Y 8HQ, England or on www.linklaters.com and such persons are either solicitors, registered foreign lawyers or European lawyers.

Please refer to www.linklaters.com/regulation for important information on our regulatory position.