March 2021

Re: General Scheme of the Online Safety & Media Regulation Bill
To: Joint Committee on Media, Tourism, Arts, Culture, Sport and the Gaeltacht
From: Twitter International Company

Introduction

Twitter shares the government's mission of making the Internet a safer place for everyone. We are an open and public platform, where people from all over the world come together for an open and free exchange of ideas. We strive to support freedom of expression, democratic debate, and healthy discourse. Indeed, our CEO, Jack Dorsey, has long said that Twitter's primary objective is to promote healthy public conversation.

Twitter commends the rigorous and highly consultative process that has produced the Online Safety & Media Regulation Bill (OSMR). Over several years, officials have studied the global regulatory landscape and consulted with a wide variety of relevant parties with an interest in this sector. The result is a comprehensive and thoughtful bill that sets a precedent at national, regional, and global levels.

This precedent will certainly be observed by countries that are currently assessing how they should legislate to meet the opportunities and challenges presented by our digital environment. It's this issue of international precedent on which Twitter would like to place particular emphasis with this submission, as we encourage the committee to consider Ireland's role in promoting legislation built on Open Internet principles.

Twitter has also contributed to Technology Ireland's submission. Therefore, the points set out below are in addition to or expanding upon Technology Ireland's paper. We would also like to point the committee to a letter that Twitter wrote with Mozilla, Automattic, and Vimeo to promote Open Internet regulation at EU level. In this letter, we talk about how the Internet is very much at a crossroads, both in Europe and around the world. The laws we pass now will shape the online world for a generation. Ireland, with this bill, is at the forefront of this movement.

Online Safety: General Comments
Twitter supports the proposed establishment of an Online Safety Commissioner within a Media Commission. We also support the collaborative development of content codes and are willing to engage in this process, drawing from our experience at EU level. Twitter also welcomes the Department's proposal to adopt a systemic approach to content regulation that takes into account the specific challenges that companies such as Twitter face.

We believe that a systemic approach is mutually preferable to more onerous and inflexible notice-and-takedown approaches where platforms and websites are incentivised to preemptively remove content to avoid liability. These regimes are ultimately counterproductive to the development of a diverse digital economy – entrenching incumbents of scale that can employ tens of thousands of moderators – and they run against the open nature of the Internet.

It will be important to ensure that any legislative proposals and subsequent codes are proportionate, technically neutral, and feasible. They should also allow for flexibility in how companies handle harmful content, given the diversity in scale and philosophy of companies within the digital ecosystem. It will be essential to consider and assess how the myriad harms manifest themselves across different platforms and jurisdictions.

Twitter recommends that any new category of "online harm" is defined clearly and precisely. The principle of legal certainty requires that those who are subject to the law should be readily able to identify their statutory obligations. Further, the challenges posed by overly vague or broad definitions are alluded to above, as it can lead to platforms taking a risk-averse approach to enforcement, thereby hastily removing content and speech that could feasibly be interpreted to come within the definition set out in laws or codes – this is known as the "chilling effect". For the avoidance of this outcome, clear definitions will be critical to remove ambiguity, helping those within scope fully understand what is required to comply with the law and, critically, ensuring that online platforms' users know exactly what is and is not permissible.

For instance, we note the inclusion of the following as an online harm: *"material which is likely to have the effect of intimidating, threatening, humiliating or persecuting a person to which it pertains and which a reasonable person would conclude was the intention of its dissemination."* First, it should be noted that much of this content would come within the scope of our existing rules, which certainly cover various forms of abuse, harassment, and hateful conduct. Second, we must acknowledge the primary importance of user intent when we think of online harms and content. Increasingly, Twitter looks to the behaviour of accounts to determine abusive intent, rather than parsing each piece of content. The latter is highly dependent on whether we have the context to determine whether a user intended to intimidate, threaten, humiliate or persecute another user(s), whereas an analysis of account behaviour reveals common patterns that are often associated with abuse and platform manipulation.

This is also the most dynamic and nebulous area of harms – often evading precise definition, with interpretations varying from case to case. It is therefore challenging to enforce at scale and with certainty of expectation between platforms, users and policymakers. If enacted, Twitter will look to work closely with the appointed Commissioner to ensure mutual agreement and precise understanding on definitions. We will also seek legal certainty on enforceability where necessary.

Further, Twitter frequently receives notices from individuals asserting that content is, in their view, illegal and must be removed. Requiring an online platform to decide what is "illegal" on the basis of a statement by one interested party, where there is no clear legal proscription of the content at issue, will necessarily result in over-censorship of legitimate content and infringement on freedom of expression. As noted

above, online platforms have neither the expertise nor the democratic legitimacy to act as arbiters of what is an "illegal" statement under the laws of a particular country and in the context of the particular factual circumstances in each case.

Twitter believes that content discussions in the coming years will, and should, focus more on content surfacing and choice rather than a binary leave-up-take-down model of moderation. Centralised moderation of content may always be with us, but the scale and complexity of online communities, and the developing technologies that underpin them, will prompt an increased focus on 'how' content came to prominence, in place of a predominant focus on whether it should be removed and how quickly. We believe the increasing focus of regulation should be on content discovery, suggestions, amplification, and propagation of content (including down- or up-ranking). Community moderation, such as that exemplified by a platform like Wikipedia, should also be encouraged because it presents a scalable and human-oriented method of maintaining a healthy and informational online discourse. Twitter is currently running its own experiment in community moderation – Birdwatch – as a way to address disinformation.

The committee should be mindful of the fact that we are already confronted with technologies that will dramatically alter the Internet e.g. blockchain. These technologies will change what we see and how we see it – some of them will decentralise services so that content is hosted across thousands if not millions of locations, presenting a potentially insoluble challenge for a centralised model of content moderation and, therefore, any regulatory system designed with only today's platforms and today's technologies in mind. Adapting our regulatory models to this horizon will future-proof them.

Global Trends & International Precedents

Twitter has a number of concerns with respect to the proposed sanctions in the General Scheme of the Online Safety & Media Regulation. Broadly, we contend that several of the sanctions as currently envisaged create a disproportionate burden on small-to-medium size companies and, in some cases, create unhelpful international precedents.

We are seeing legislative models being taken from one country and copied multiple times across other countries. Where those models protect the Open Internet, this legislative exporting is a positive development. However, we see a worrying trend whereby a country with strong due process protections will enact a particular law with stringent sanctions and this law is then copied by another country that may have weaker protections. In this way, more repressive regimes can exert pressure on content platforms to comply with requests that would restrict freedom of expression, journalism, and activism. This is a particular concern for Twitter, which is uniquely associated with such forms of speech.

We, therefore, urge policy-makers to be mindful of their unique role at this time. The OSMR is breaking new ground, largely for the positive. However, some of the proposed sanctions will surely be observed by other countries and may be adopted to pursue more oppressive ends. Already, we see countries imposing punitive financial penalties to make the business environment difficult for platforms with unwelcome views on freedom of expression. Throttling and blocking is used to limit citizen access to news, information, and minority or opposition perspectives. And non-compliance is met with harassment and custodial threats for company directors. This digital oppression by nation states is gaining momentum every week in every corner of the world. It is challenging the foundational principles of a global, free and open Internet. We humbly ask that Ireland acknowledge this trend and reflect its support for the Open Internet in the sanction regime set out by the OSMR.

More specifically, we would encourage the removal of the director liability from the OSMR for the sake of consistency with the chosen systemic framework. With the focus of oversight directed at how

organisations are meeting certain standards, a penalty system that sanctions the organisation, rather than the individual, provides an intuitive and effective penalty regime.

Equitable Sanctions

The General Scheme states that companies may be liable for financial sanctions of “up to €20,000,000 or, in the case of an undertaking, up to 10% of relevant turnover of the preceding financial year, whichever is higher.” First, it would be helpful to further clarify whether turnover is calculated on a national or international basis.

Second, while it’s acknowledged that financial sanctions would be measured in a manner proportional to breaches, it should also be recognised that there is a large cohort of platforms for which financial sanctions on the order of those set out in the OSMR would be an existential concern. The technology industry is not monolithic – there is significant divergence in philosophical approach, in revenue generation, and in corporate footprint and resourcing. This echoes a point from the introduction to this document: Twitter, while large in impact, is not among the very largest or wealthiest companies in the technology sector. For example, our global employee headcount is less than half of what some companies have in Ireland alone.

For illustrative purposes, the pie chart in this Tweet highlights the gap between the largest companies and “everyone else”. In our advocacy for regulation that preserves the Open Internet, we often talk about how laws that take a whole-of-industry view should also promote fair competition, rather than entrenching the advantages of incumbents. The implementation of equitable sanction regimes are a key part of promoting such fair competition, thereby ensuring that the platform sector remains open to new businesses, new ideas and entrepreneurs who require some measure of assurance that their enterprise can co-exist in the same space with the giants of our industry.

Regional Alignment

Twitter supports regional and global regulatory alignment around systemic models insofar as it’s possible. A coherent national and regional approach to content regulation provides the clarity that cross-border services require to fulfil user expectation that their experience of our platform will be consistent regardless of where they are in the world. At the core of this is the preservation of the Country of Origin principle.

In this regard, we have been witnessing trends to regulate online platforms at a national level by other EU State Members and we are concerned that this fragmented approach will complicate any possibility of a more coherent and pan-European regulatory response. Indeed, we believe that the free and open Internet needs a coherent set of standards at European level. Otherwise, any new regulation would risk building virtual walls between our digital communities.

Regional alignment of regulation is also important for small-to-mid size tech companies. In fact, it may also be existential. Were there to be separate and distinct regulatory regimes across Europe, and the world, it may incentivise a highly risk-averse compliance model that would change the nature of our and other services. For some companies – those existing and those yet to be founded – it would essentially put a hard ceiling on growth, as the compliance burden created by divergent standards would be too onerous on resources. Again, this highlights the function of regulation as a tool for encouraging fair competition between companies of all sizes and capacities.

Diverging regulations would also have the fundamental effect of fragmenting the Internet. Broadly speaking, a few exceptions notwithstanding, the understanding of what the Internet is and what it offers is

universal – a user's experience of Twitter in Ireland should, ideally, be no different than anywhere else. Regional and global alignment on regulation would preserve this understanding for citizens and preserve the defining traits that have made the Internet an unprecedented catalyst for economic, educational and cultural progress.

We believe that Ireland has a unique role to play in advocating for this regulatory cohesiveness. Its own proposals are evidence-based and follow developing international best practice. Ireland can become a leader in convening like-minded countries and advocating for a systemic regulatory approach that will lay the foundation in Europe for the next twenty years of the Internet – protecting consumers, encouraging openness and freedom, while promoting business innovation.

Reconciling Existing Laws & Proposals

Recently, the EU's Copyright Directive has entered into force and the draft Digital Services Act ("DSA") was published by the European Commission. The AVMS Directive, the new Copyright Directive and the proposed DSA lay down rules for online platforms. It is understood that the provisions in the AVMS Directive and in the Copyright Directive are based upon and complement the DSA. Nevertheless, the interaction between the directives is not fully clear. For example, the Copyright Directive largely places the responsibility to prevent the availability of protected works on platforms themselves, while the AVMS Directive not only involves platforms, but also the uploaders and users that exercise influence on their service. These contradictions create legal uncertainty and are harmful to small and medium-sized companies, as well as to users' freedom of expression. More broadly, there is a need to reconcile and perhaps rationalise the wide range of overlapping definitions set out under the EU Directives and national laws.

Funding

Twitter notes the proposed funding model for the Media Commission and is supportive of the proportionate measurement of contributions based on regulable activity in Ireland. In this context, we note that Twitter is still exploring all available options under the Audiovisual Media Services Directive and we have not yet determined whether the new rules on VSPs apply to our service.