CENTRE FOR MEDIA TRANSITION

Exposure Draft of the Social Media (Anti-Trolling) Bill 2021, December 2021

Submission to Attorney-General's Department

DATE: 21 January 2022

About the Centre for Media Transition

The Centre for Media Transition (CMT) is an applied research unit based at the University of Technology Sydney (UTS). Launched in 2017, the CMT is an interdisciplinary initiative of the Faculty of Arts and Social Sciences and the Faculty of Law.

CMT works across disciplines to explore and develop responses to the dramatic and ongoing movements wrought by digital disruption to the media industry, the role of journalism in Australian democracy and the world more widely, and the business models that support a diverse and prosperous industry.

This submission was prepared by:

- Dr Derek Wilding
- Dr Sacha Molitorisz.

Research assistance by Vidya Kathirgamalingam.

Contact

Centre for Media Transition
Faculty of Law, University of Technology Sydney
Building 2, Level 15
UTS City Campus, Broadway
PO Box 123, Broadway NSW 2007

cmt@uts.edu.au
+61 2 9514 9669

cmt.uts.edu.au

Introduction

Thank you for the opportunity to make this submission on the Exposure Draft.

We recognise that other researchers and commentators have expressed views about perceived shortcomings with this draft Bill. While we support some of these points, in this submission we address what we believe to be the significant progress the Bill makes in encouraging social media providers to implement a responsive complaint handling scheme.

In making these comments, we draw on three separate but related research activities of the Centre for Media Transition:

- Our work on defamation law reform, including our submissions to the earlier discussion papers issued by the Council of Attorneys-General;[1]
- Our work on regulation of digital platforms, including our research report for the ACCC's Digital Platforms Inquiry; written submissions made to that inquiry; and articles published in the *Journal of Telecommunications and the Digital Economy* and the *Journal of Media Law*;[2]
- Current research on digital platform complaint handling, funded by a research grant from the Australian Communications Consumers Action Network;
- Current research on the future of news media standards councils.

1. The need for a dispute resolution scheme

As the Explanatory Notes make clear, there is a demonstrated need for schemes that help users of digital platforms take action in response to harmful material about them posted on social media. Social media providers have themselves made progress in establishing rules about the type of content that can be harmful and in providing customers with ways of flagging such content. Digital platforms have recently established a complaints portal and independent complaints committee as part of their commitments under the Australian Code of Practice on Disinformation and Misinformation. This initiative is worthwhile and is to be encouraged, but there are two ways in which it is limited in scope. First – and this is not a criticism of the initiative itself, but an observation on scope – it is limited to material that can be regarded as 'disinformation' or 'misinformation' and does not extend to other types of content. Second, complaints under this scheme concern the platforms' compliance or otherwise with their own obligations under the code; consumers cannot make complaints to that committee or and seek remedies in respect of specific online content they encounter.

These limitations mean there is a need for further mechanisms to address other problematic content. In our view, if platforms have not established such mechanisms themselves, it is reasonable for government to step in and provide the framework within which these problems can be addressed.

---

[1] Policy submissions are listed under Defamation Law Reform. See also, Centre for Media Transition (2018), *Trends in Digital Defamation: Defendants, Plaintiffs and Platforms*.

[2] Wilding et al (2018), *The Impact of Digital Platforms on News and Journalistic Content*; Wilding, Derek (2021), 'Regulating News and Disinformation on Digital Platforms: Self-Regulation or Prevarication?' *Journal of Telecommunications and the Digital Economy* 9(2) 11-46; Lee, Karen and Molitorisz, Sacha (2021), 'The Australian News Media Bargaining Code: Lessons for the UK, EU and Beyond' *Journal of Media Law* 13(1) 36-53.

2. The scheme proposed in the draft Bill

The motivation for a social media service provider to establish the kind of complaints scheme anticipated by the Bill arises out of the imposition of liability (under s 14) as a publisher of third party content that could be defamatory. The Bill appears to clarify this aspect of the position at common law, making social media services liable alongside the author of the comments, but it goes further in establishing that a defence of innocent dissemination is not available to the social media service. Instead, there is a new defence (in s 15) that can be used by the social media service where it has established and followed a complaints scheme that, in effect, puts the complainant in direct contact with the author of comments that might have been made anonymously or pseudonymously. The scheme thereby gives the complainant the opportunity to resolve the matter informally or to sue the commenter for defamation.

3. Limitation to defamation

There is a vast range of online content that is in some way harmful or otherwise unacceptably offensive, and there are now multiple schemes in effect or in development to try to address these different forms of content. Some of these overlap with defamation. When developing its own approach to defamatory online content, the Law Commission of Ontario made the following observation:

> Trolling, flaming and cyberbullying can all involve defamation … There is increasing concern about the ease with which online reputation can be harmed and the lack of any checks and balances on the torrent of offensive content that is publicly and permanently available online. Although not always labelled as such, much of this content may be defamatory.[3]

The fact that the scheme in this draft Bill is limited to defamatory content rather than other forms of content, including the more expansive category of 'trolling', is underlined by the operation of s 16(1)(h), which establishes that the social media service need not take any action in response to a complaint if it reasonably believes that the request 'does not genuinely relate to the potential institution by the complainant of a defamation proceeding against the commenter'. This means that in order to filter out complaints that are about something other than defamation, some consideration based on a knowledge of defamation law will need to be applied at the first stage of the complaints scheme.

There are two risks that we think arise from the way in which the Bill frames the complaints scheme. First, there is a risk that the Bill would channel more disputes about online content into the court system to be addressed as defamation complaints, when it would be preferable to have them resolved via a more informal mechanism. In fact, the scheme might even incentivise platforms to facilitate the removal of these matters into the court system in order to avoid any further role in their resolution. Second, there is a risk that complaints that do not meet the requirements at common law for defamation, but which are potentially more significant than defamation complaints while not amounting to cyber abuse or the other matters covered by the new *Online Safety Act 2021* (Cth) would slip through the gaps in this framework.

---

[3] Law Commission of Ontario (2020). *Defamation in the Internet Age: Final Report*, p.4.

We think the Bill represents a good first attempt to come to grips with this problem, but we think it would be worth considering how the complaints scheme might be adapted in two ways – firstly, to address matters other than defamation, and secondly, to push these matters into a more comprehensive alternative dispute resolution scheme. We expand on these points in our final section below.

4. Fair treatment

We have some concerns about the way the various participants in the chain of publication are treated under the Bill. From our reading, it appears that:

- Liability of authors of defamatory comments remains the same
- Liability of page owners is removed entirely
- Liability of social media services is expanded.

Our understanding of these arrangements is based on the combined effect of ss 14, 15 and 16, which establish who will be liable as a publisher of third party comments and the circumstances in which certain defences will be available.

A complaints scheme must comply with certain ‘prescribed requirements’ (set out in s 16), including notifying the commenter of the complaint and, in effect, providing an opportunity for the commenter to agree to have the comment taken down. Where the social media service meets the prescribed requirements and follows the procedure but the complainant is ‘dissatisfied’ with the outcome of this initial notification and opportunity for removal, the social media service may be prevented from relying on the new defence. This will happen where the complainant requests the contact details of the commenter but the social media service does not provide these details to the complainant, including in circumstances where they hold those contact details but the commenter refuses the request for them to be passed on to the complainant.

The Bill also reverses the recent common law clarification (by the High Court in the Voller decision)[4] of the liability of page owners. The Voller decision was concerned with the conduct of page owners that are news media publishers; it established that news publishers who post news reports on their Facebook pages will be liable for any subsequent comments posted by third party users. The Bill makes the necessary adjustment to this position, relieving publishers of liability for the comments of others. It also extends this immunity for third party comments to *all* page owners.

The effect of these provisions would be:

- Page owners, including news publishers, other businesses and purely social sites established by anyone in the community, will be immune from liability for comments posted by others, no matter how careless or complicit the page owner is in providing the setting for the comments to be made.
- Even where they have established and followed a complaints scheme that complies with the Bill, social media service providers will be liable for third party comments unless the commenter agrees to share their contact details.

---

[4] *Fairfax Media Publications Pty Ltd v Voller; Nationwide News Pty Limited v Voller; Australian News Channel Pty Ltd v Voller* [2021] HCA 27.

We note that the Bill does not affect the availability of other defences, such as the common law defence of truth or of justification under s 25 of the Uniform Defamation Law (UDL), but the difficulties of establishing the other defences are well known and will be greater for a party that has no direct connection with the author of the material.

Although s 14(3) of the Bill appears to greatly expand the liability of social media services by removing the availability of the defence of innocent dissemination, the new 'complaints scheme defence' can perhaps be seen as a partial replacement of innocent dissemination. The concept of innocent dissemination is something that, whether at common law or in legislation, and whether as part of the UDL or some other scheme, needs to be preserved. In the context of third party comments on social media, it is our view that a social media service – not being the originator of the comments – should not be liable for them if it was not put on notice of the defamatory nature of the comments or if, once on notice, it took appropriate action within a reasonable period. To some extent, the Bill does take this approach. Although the social media service won't have access to the established defence of innocent dissemination at common law or under s 32 of the UDL, it will have a replacement defence if it establishes a compliant scheme under the Bill and follows the procedures, including giving the commenter's contact details to the complainant when requested. Further, it appears that if a social media service has a compliant complaints scheme in place and a complainant *does not use that scheme* – so that the social media service is not put on notice – the defence will still be available.

For the record, if our understanding of this is incorrect and the defence won't be available in these circumstances, we think the Bill should be amended so it does work in this way. But if it does work this way, then the expansion of liability on the part of social media services, and the imbalance among the various participants in the chain of publication, is not so great as appears on a first reading. In our view, if our understanding is correct, this is an appropriate apportionment of liability

The remaining problem we see with the new defence concerns the situation where the commenter does not agree either to remove the comment (which might, in many circumstances, resolve the problem) or to have their contact details passed on to the complainant. We understand why there might be a reluctance to provide a complete defence at this stage if all the platform does is pass on the complaint and nothing more comes out of the complaints procedure, but we also think it is a very harsh outcome for the social media provider's defence to be wholly contingent on the cooperation of the commenter – even where the provider has implemented an effective process for establishing contact details and for facilitating notification of disputes, and possibly even where the provider has removed the material under its terms of service regardless of the agreement of the commenter.

Our suggestion is not to abandon the scheme proposed in the Bill, but to build on it by having the platform go one step further in facilitating a resolution. Our proposal is explained further below. We think that if social media services do take this additional step, they should have the benefit of the 'complaints scheme defence' as a replacement for the innocent dissemination defence, even if this doesn't give the complainant the outcome they're seeking.

5. Recommendation: facilitating resolutions

The draft Bill represents a significant advance on past approaches to dealing with complaints about social media content. It starts the process of having platforms become active participants in

resolving problems about content they did not create or curate but for which, as distributors, they ought to have some responsibility. In submissions to past reviews, we have consistently put the position that digital platforms should not be regarded as 'publishers'. This is because we think it's desirable to retain the coherence that has attached to that term in the past and which will continue to be important as a justification for a range of obligations (such as accuracy and fairness in news reporting) as well as entitlements (such as financial support and relief from certain obligations including privacy protections in the course of journalism).

For this reason, our first preference is to see the development of social media complaints schemes decoupled from liability under defamation law. We note that in its recommendations for reform, the Law Commission of Ontario (LCO) advanced a scheme that required social media companies to participate in complaint handling, but did not attribute to them the status of publishers. Instead, a provider might be liable for a statutory penalty for failure to comply with the scheme. Further, the procedures supported by s 5 of the UK Defamation Act mean that liability only runs from the receipt of a notice from someone who isn't able to identify the commenter. However, we appreciate that the draft Bill is the first real attempt to tackle the problem in Australia, and we see value in continuing to develop these ideas, wherever they may ultimately reside. Accordingly, we offer the following observations for consideration as part of this scheme or a related scheme.

1. A significant potential benefit of this scheme is that the act of notification from the social media provider to the commenter could itself prompt a commenter to remove or modify a comment, thereby avoiding the need for litigation. We think it would be worth exploring how this could be taken further through **a second stage of dispute resolution**. This approach would build on that set out in the draft Bill and move further towards the UK scheme and that proposed by the LCO. Both of these international approaches involve a further attempt to create a forum for the resolution of complaints between the complainant and the commenter, without the platform exercising judgement over the defamatory nature of the content. The LCO describes this as 'no assessment of merits', with the platform essentially fulfilling the function of a 'go-between' (p.91).
2. In order to help businesses and community members resolve matters that might or might not constitute defamation – or which constitute defamation but are not so serious as to justify the vast costs and small damages that can result from social media defamation claims – **the alternative disputes scheme could extend to matters other than defamation** and to content other than third party comments. One of the criticisms made of the draft Bill is the potential loss of anonymity in circumstances where revealing someone's identity might itself have harmful consequences. While we think it should be possible to account for this in limited circumstances and to require the protection of personal information by social media providers (and we note that the Ontario proposal maintains anonymity through an interim take-down and put-back stage), we think the loss of anonymity could be countered by a gain in making the scheme available beyond complaints about damage to reputation. This could include, for example, disputes over content that is said to offend against community standards but that does not amount to Class 1 or Class 2 content under the new online content scheme under the Online Safety Act.
3. While we encourage the Department to consider how the Australian scheme could move closer to the UK or Ontario model, an alternative – which may take the resolution process a step further while still avoiding social media providers making judgements on the content of these disputes – would be to **involve an independent body**. Regulation in this area is difficult because technologies, business practices and consumption habits evolve rapidly and because regulation needs to protect freedom of speech at the same time as addressing online harms.

This may be a reason for using self- or co-regulation in preference to direct government regulation, but in-house schemes that are developed by and apply to a single company may not adequately address community standards and user expectations. They may also fail to attract sufficient confidence within the community. An independent industry body, properly constituted to transparently apply established rules, may be able to provide more significant outcomes for complainants. As we have pointed out in a series of submissions, regulation of content online requires an approach that is as coherent and holistic as possible, so that issues such as defamation, trolling, free speech, and more, are regarded and dealt with as interconnected rather than unrelated.

4. If an approach involving an independent body is considered, some threshold issued will need to be addressed. Given the vast amount of content being posted on social media, scale is potentially a problem in online content disputes, so some filtering mechanism is needed. While some aspects of complaint handling might be performed as part of an automated Online Dispute Resolution (ODR) system, it is not clear yet whether automated ODR represents a viable alternative to systems that involve human agents.
5. The issue of who might handle the complaints that ultimately require some human intervention remains to be considered. As part of the separate research activity mentioned above, CMT is currently looking at the ACCC's recommendations 22 and 23 made as part of the Final Report of the Digital Platforms Inquiry. We expect to release a report on this work in the coming months. At this stage we note only that the subject matter of the current draft Bill, being concerned with content complaints, appears to be somewhat removed from the interests and expertise of a body such as the Telecommunications Industry Ombudsman. It fits more closely with the activities of a body like the Australian Press Council, and the involvement of press complaints schemes with resolution of defamation claims has precedent in that in the UK both IPSO and IMPRESS can refer matters to an arbitration scheme for resolution. We have noted previously that we think there is need for reform of the news standards schemes in Australia and that cross-platform production and distribution of news call for a cross-platform standards scheme. We have also suggested that digital platforms could be brought into an overhauled standards scheme in order to support trusted news sources without attracting the same content obligations (such as accuracy and fairness) as publishers. If that were to happen, it may be possible for the **expanded and overhauled media standards body to provide the home for escalated complaint-handling under the social media complaints scheme** discussed here. Even if this were not possible, an industry-based scheme could provide consumers with information on how to make and where to direct their complaints. Under the scheme proposed in the draft Bill, there is a need for information on what constitutes defamation. This applies to complainants but also to commenters who are asked to make a decision on whether to take down content. While defamation law is complex, providing this kind of information is not impossible. At UTS, we recently developed a one-to-two hour online course on defamation, taking account of the recent reforms.[5] A streamlined version of a module such as this could be made available for potential complainants and respondents.

[5] See: Defamation and Digital Publishers: How to Navigate Legal Risk (on-demand).