RESPONSES TO DISCUSSION PAPER

ATTORNEYS - GENERAL

REVIEW OF MODEL DEFAMATION PROVISIONS – STAGE 2

SUBMITTED BY STUDENTS ENROLLED

IN LAWS5084 – MEDIA LAW 2021[1]

19 May 2021

These responses conform to the numbering of consultation questions within the Discussion Paper

[1] These responses contain the work of law students in group projects enrolled in Media Law at Macquarie University and this document and its contents do not reflect the position of Macquarie University which takes no position on these matters

EXECUTIVE SUMMARY

Question 5: Treatment of Internet Intermediaries as Publishers of Third-Party Content

(a) Should internet intermediaries be treated the same as any other publisher for third-party content under defamation law?

No - Internet intermediaries should not be treated in the same manner as any other publisher for third-party content under Australia's defamation laws but, under some circumstances they should incur liability for defamatory materials on their websites.

Internet intermediaries typically lack the level of control over third-party content enjoyed by more traditional publishers and the volume of material frequently far exceeds the amount generated in traditional print or broadcast media.

Internet intermediaries should incur liability for third-party content in two situations: (1) where the intermediary actively edits or otherwise participates in the content or (2) where the potentially defamatory materials are brought to the intermediary's attention and the materials are not promptly removed from the website. The intermediary should not be held liable for erroneous removal of lawful material.

Question 6: Immunity for Basic Internet Services

It is necessary to provide immunity from liability in defamation to basic internet services. However, basic service providers should only enjoy immunity where they are acting as mere facilitators to internet services and not acting beyond this capacity. Basic internet service providers only offer access to the internet and have no active

contribution to the generation of content. It is unreasonable to suggest basic service providers should monitor all websites existing on the internet for potentially defamatory content. If they were to be held responsible for all defamatory content passing on the network, then their continued commercial viability would be doubtful.

Question 7; Amend Part 3 of the MDPs to Better Accommodate Complaints to Internet Intermediaries

Some amendments are necessary because Part 3 does not mention internet intermediaries. This needs to be clarified so that aggrieved persons would be able to give notice to 'direct or indirect publishers or authorisers.' Further, the notice would need to provide a specific link to the offending material and list all of the parties to whom the notice was given.

Question 8: Clarifying the Innocent Dissemination Defence

A standalone section (clause 32A) should be enacted which creates an innocent dissemination defence for digital platforms and forum administrators. The nature of social media and the internet means that the current 'subordinate distributor' definition is not appropriate for digital platforms and forum administrators. However, innocent dissemination is not an absolute defence and if platforms promote defamatory content then the defence fails.

Question 9: Safe harbour Subject to a Complaints Notice Process

A defence akin to that of section 5 of the Defamation Act 2013 (UK) should be included, in at least some capacity, into the Model Defamation Provisions. The legislation in its present form is already drafted in a manner suitable for Australia The defence is justified in

that it does not punish operators for communications they did not publish.

Question 10: Immunity for Internet Intermediaries Unless They Materially Contribute to the Unlawfulness of the Publication

Yes, immunity should be the rule. Fundamentally, internet intermediaries are not the primary publishers of content. Currently the basis on which these third-party internet intermediaries are liable for the actions of individuals users is confusing and largely incoherent. The main limiting devices of liability, mainly intention, passivity and knowledge, are ineffective in drawing a clear distinction for circumstances in which intermediaries will not be held liable.[2] An immunity would bring greater clarity whilst still holding intermediaries liable when they have taken responsibility over the content. Or,

No, blanket immunity should not be provided to digital platforms for third-party content even if they do not materially contribute to the publication. Immunity should be provided to organisations if they remove content in a reasonable time once they have been notified of an unlawful publication. If organisations are notified of the content and take an unreasonable amount of time to remove the material, blanket immunity should not be extended. Even if organisations do not materially contribute to publications, by keeping unlawful content available on their platform, they are contributing to the publication by gaining an audience for the material. By allowing the unlawful publication to remain on their website following notification, the digital platform is knowingly assisting in the publication of unlawful

[2] Kylie Pappalardo and Nicolas Suzor, 'The Liability of Australian Online Intermediaries' (2018) 40(4) *Sydney Law Review* 469.

material. Due to this, blanket immunity should not be extended once notification has been provided.

Question 17: Other Issues Regarding Liability of Internet Intermediaries

There were a variety of proposals that do not lend to generalisation. However, one of special note relates to the concept of imposing special liability where prohibited content relates to protected persons. Special liability on internet intermediaries would be imposed upon failure to promptly remove third-party prohibited content involving protected persons

a. These persons may include:
    i. People with physical or mental handicaps;
    ii. Indigenous Australians;
    iii. Elderly people; and
    iv. Transgender and intersex people.
b. Increased damages would be available under the statute for content involving the above-protected class.

End of Executive Summary

Detailed Responses Follow

Question 5: Treatment of Internet Intermediaries as Publishers of Third-Party Content

(a) Should internet intermediaries be treated the same as any other publisher for third-party content under defamation law?

No - Internet intermediaries should not be treated in the same manner as any other publisher for third-party content under Australia's defamation laws but, under some circumstances they should incur liability for defamatory materials on their websites.

Internet intermediaries typically lack the level of control over third-party content enjoyed by more traditional publishers and the volume of material frequently far exceeds the amount generated in traditional print or broadcast media.

Internet intermediaries should incur liability for third-party content in two situations: (1) where the intermediary actively edits or otherwise participates in the content or (2) where the potentially defamatory materials are brought to the intermediary's attention and the materials are not promptly removed from the website. The intermediary should not be held liable for erroneous removal of lawful material.

While some of the contributors would impute liability to Internet Service Providers under limited circumstances, the clear majority concluded that ISPs as providers of basic internet service act as mere conduits and should not face liability for content sent over the network. However, the situation should be considerably different for the intermediaries that actively host third-party content like Facebook. For the purposes of the Model Defamation Provisions (MDPs), pursuant to the *Broadcasting Services Act,* internet service providers ('ISPs') are 'mere conduits' and not liable.[3]

[3] *Broadcast Services Act 1992* (Cth) sch 5 cl 91.

A person is a publisher for the purpose of the MDPs if the publication in question is, or may be, defamatory of another person.[4] Through the 2020 decision in *Fairfax Media v Voller*, internet intermediaries are now treated as publishers.[5] Publication is the process of communicating the matter in a comprehensible form to a third party.[6] Internet intermediaries are liable for publication,[7] republication,[8] multiple publication,[9] and even omitting to facilitate the removal of third party defamatory content.[10] However, in imposing these obligations, the law has failed to contemplate the immense volume of third-party content on intermediaries. For example, in 2020, Facebook had an estimated 11.23 million users in Australia alone,[11]and approximately 2.7 billion active users worldwide.[12] With the sheer number of people using Facebook as a tool for communication, the challenges for Facebook to remove all defamatory content as a publisher are self-evident.

The global span of many intermediaries means that it is both impractical and misguided to treat them as traditional publishers.[13] The MDPs, as they currently stand, will struggle to overcome this challenge. Intermediaries have given all individuals the tools to become a publisher. While intermediaries should not be treated as publishers, they should not escape liability altogether. It is inappropriate to treat non-ISP intermediaries as 'mere conduits'. Individuals utilize intermediary platforms to publish, with the knowledge that engaging with the medium promotes further dissemination of the content.[14]

---

[4] *Defamation Act 2005* (NSW) s 12.
[5] *Fairfax Media Publications; Nationwide News Pty Ltd; Australian News Channel Pty Ltd v Voller* (2020) 380 ALR 700 ('*Voller*') (review granted)
[6] David Rolph et al, *Media Law: Cases, Materials and Commentary* (Oxford, 2nd ed, 2015) [7.4.1].
[7] *Voller* (n 5).
[8] *Sims v Wran* [1984] 1 NSWLR 317.
[9] *Dow Jones & Co Inc v Gutnik* (2002) 210 CLR 575.
[10] *Byrne v Dean* [1937] 1 KB 818.
[11] Thomas Hinton, 'Number of Facebook Users in Australia from 2015 to 2022' (Web Page, Statista, 2021) <https://www.statista.com/statistics/304862/number-of-facebook-users-in-australia/>.
[12] H Tankovska, 'Number of Monthly Active Facebook Users Worldwide as of 4th Quarter 2020' (Web Page, Statista, 2021) < 'Number of Monthly Active Users https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/>.
[13] See Marshall McLuhan, *The Gutenberg Galaxy* (University of Toronto Press, 1962).
[14] See Marshall McLuhan, *Understanding Media* (McGraw-Hill, 1964).

It is more appropriate for a new category of liability to be imposed on the MDPs, superseding the decision in *Voller*,[15] to balance practicality and responsibility. We propose a new classification as an 'active conduit.' This category recognises that intermediaries have a role in aiding dissemination, but do not publish content per se. Liability is triggered by receiving information that content may be defamatory.

Possible Solutions:

The defamation laws should provide that once an internet intermediary is made aware of possible defamatory content, it becomes an 'active conduit' of the offending material. Each intermediary must provide a user-friendly and accessible reporting mechanism by means of which offending content may be reported. Then an internal review must be undertaken by the intermediary to either remove or retain the matter and then provide some reason for the decision. Having become an 'active conduit' of the information, then the intermediary must assume the same defamation liability as any other publisher. To avoid liability as a publisher, the intermediary must act within forty-eight (48) hours after receiving the notification. The law must further provide that internet intermediaries do not bear liability for erroneously removing lawful third-party content.

As an alternative solution, an independent authority may be constituted to review allegedly defamatory content. Intermediaries should be responsible to remove content based on *legally binding recommendations* from this body. A drawback to this approach is who will bear the cost of this independent review and will such a body be able to act promptly?

[15] *Voller* (n 5).

Question 6: Immunity for Basic Internet Services

(a) Is it necessary and appropriate to provide immunity from liability in defamation to basic internet services.

It is necessary to provide immunity from liability in defamation to basic internet services. This is because of the fundamental role they play in facilitating access to internet services. Accordingly, basic service providers should only enjoy immunity where they are acting as mere facilitators to internet services and not acting beyond this capacity. Basic internet service providers only offer access to the internet and have no active contribution to the generation of content. It is unreasonable to suggest basic service providers should monitor all websites existing on the internet for potentially defamatory content. If they were to be held responsible for all defamatory content passing on the network, then continued commercial viability would be doubtful.

As basic internet services providers are merely acting as content neutral vehicles for communications, they should be granted indemnity. If liability were to extend to basic internet services, it would be a logistical impossibility for them to be able to monitor everything passing through their system. Further, it would be impractical for Australia to adopt such a position alone as defamation laws vary among nations.

(b) If such an immunity were to be introduced, should it be principles-based or should it specifically refer to the functions of basic internet services?

There were contrasting views, with one of the student groups favouring a principles based and the others preferring others a function based approach.

Principles based Approach –

The immunity should be implemented on a basis of principles as opposed to specifically referring to functions of basic internet services. It is proposed that the implementation of principles-based immunity would permit the law to keep up with constantly evolving technology and changing functions of basic internet services. The definition of 'basic services' may be somewhat subjective and as a result, focussing on principles would allow for clearer statutory interpretation.

Functions based Approach –

An immunity for basic internet services from liability for defamation should be centred around its functions; namely, passivity and neutrality. If a basic internet service that is ordinarily not involved in the creation and dissemination of content provides an avenue for the publication of defamatory content (such as a discussion forum) for the benefit of its customers, they should not be held liable for content published on that platform. Of course, an important qualification of this immunity would be 'actual awareness' (*Discussion Paper)*, of the defamatory content.

This immunity is based on the idea that when a basic internet service provides a forum for discussion, that platform is an extension of their neutral and passive function of facilitating access to the internet and ensuring all customers are satisfied with their services.

It is worth noting that we advise against a principles-based immunity for basic internet services due to the difficulties associated with formulating it. Namely, whether to take a narrow or broad approach to the immunity, as well as what to incorporate into the provisions. We do however acknowledge the difficulties surrounding the broad scope of our proposed functions-based immunity.

(c) Are there any internet intermediary functions that are likely to fall within the definition of basic internet services that should not have immunity?

Undoubtedly there exist some content-neutral functions of internet intermediaries which might fall within the definition of basic internet services. If this is the case, taking a functional approach, then the intermediary should be able to take advantage of the immunity offered to internet service providers. But as a countervailing factor, the level of content moderation offered by the entity in question should also be considered. Intermediaries involved only in redirecting/referring users to external content such as search engines may be granted greater degrees of immunity than most intermediaries.

(d) Is there a risk that providing a broad immunity to basic internet services would unfairly deny complainants a remedy for damage to their reputation?

A broadly defined immunity for basic internet services would naturally increase the difficulty of accessing a remedy for defamation because it would diminish the number of potential defendants available to the complainant. But diminishing the number of potential defendants is not necessarily unfair. Mitigating the risk of being *unfairly* denied access to a remedy therefore, is a matter of establishing the appropriate scope of the immunity with sufficient clarity, such that the plaintiff would not be barred from bringing a claim against defendants that ought to incur liability in respect of a defamatory expression. Presumably, potential defendants would include not only the author of the material but also intermediaries who were not basic internet services.

Question 7; Amend Part 3 of the MDPs to Better Accommodate Complaints to Internet Intermediaries

(a) How can the concerns notice and offer to make amends process be better adapted to respond to internet intermediary liability for the publication of third-party content?

Some amendments are necessary because Part 3 does not mention internet intermediaries. This needs to be clarified so that aggrieved persons would be able to give notice to 'direct or indirect publishers or authorisers.' Further, the notice would need to provide the specific link to the offending material and list all the parties to whom the notice was given.

By listing all the potential parties in the notice, the defendants would have an opportunity to craft an appropriate settlement drawing upon their combined resources. This could help in minimising litigation.

(b) What are the barriers in the concerns notice and offer to make amends process contained in Part 3 of the MDPs (as amended) that prevent complainants from finding resolutions with internet intermediaries when they have been defamed by a third-party using their service?

It is sometimes difficult to trace and identify the author of defamatory material on the internet. There are thousands of posts shared on the internet. Further, from the intermediary's perspective, it is difficult to track and be aware of each post and each sharing of the originating posts.

This notice process is ill-suited to the tech era; there are too many platforms and people responsible - particularly with the oversaturation of platforms available to re-post, re-share and provide links for the material in question. Material continues to be reproduced and re-shared on other platforms and saved onto computers and phones (screenshots).

There is not a mandatory timing requirement. An Intermediary is only required to make an offer 'as soon as practicable' as per clause 18(1)(b). This

provides for uncertainty to the aggrieved person. Further clarification on the timeline could therefore be useful to increase transparency.

Currently there is no requirement for the complainant to provide the URL of the defamatory material. Adding such a requirement would allow for a specific location to be identified which could be used to notify the publisher. On the other hand, including an URL does not always assist in finding the originating source, e.g. when URLs do not remain uniform.

(c) In the event, the offer to make amends process is to be amended, what are the appropriate remedies internet intermediaries can offer to complainants when they have been defamed by third parties online?

Initially, removal of defamatory content from the website and removal of the content/links from search results (if the internet intermediary is a search engine provider). However, this raises the question of whether internet intermediaries have appropriate authority to remove with legal impunity the 'defamatory' content on their platforms without the third-party's consent.

Beyond this initial step, current remedies related to the offer to make amends process may not be suitable when applied to internet intermediaries:

- Apologies may be offered but they are potentially not an appropriate remedy. An issue would include identifying the conduct they are apologising for.
- Reasonable corrections may not be appropriate. This is due to difficulty in tracking the defamatory material and correcting every post/repost/share/republication that is available on the intermediary's platform.

Possible alternatives include:

- Transparency from the intermediary in what they are doing to investigate and respond to the content/issue:

    - Establish a general contact such as a Regulatory Compliance Unit so that people can identify who is responsible for deciding their cases.
    - Provide monthly reporting outlining current practices and processes in relation to responding to defamatory material.
- Potential introduction of a badge on content, such as the badge introduced by Facebook that identifies material as being potentially false or having incorrect information in it (e.g. fake news). Such a badge could include a response/reply from the person who alleged the information was defamatory or notification that the information may be inaccurate or is being investigated.

Question 8: Clarifying the Innocent Dissemination Defence

This response subsumes parts (a) through (e) of Question 8

A standalone section (clause 32A) should be enacted which creates an innocent dissemination defence for digital platforms and forum administrators. The nature of social media and the internet means that the current 'subordinate distributor' definition is not appropriate for digital platforms and forum administrators. However, innocent dissemination is not an absolute defence and if platforms promote defamatory content, then the defence fails.

CURRENT CLIMATE

Currently, social media platforms are seen as a primary distributor under the *Defamation Act*. This is an issue as it puts the onus on the platform when the individual is in fact responsible for posting the content online. There is also too much content on social media for the platforms to effectively monitor it in a timely manner. It may be difficult for a social media platform to determine which content is actually defamatory, particularly when considering the various defences that may be available in different nations. The individual who is posting the content is likely

to be more informed and have greater insight surrounding the post in question. Whether social media platforms should have the authority to monitor matters related to defamation also is an issue of contention. It is important not to incentivise the over-removal of content. Such a practice impedes legitimate freedom of speech.

FORUM ADMINISTRATOR AND DIGITAL PLATFORMS

In response to current issues, the innocent dissemination defence should be extended to encompass digital platforms and forum administrators. A standalone section for digital platform and forum administrator would be a preferred alternative, as the nature of social media and the internet cannot be compared with the current 'subordinate distributor' definition. A separate provision will allow for legislators to adapt this defence without impacting the current and successful s 32 defence as technology evolves. Proposed language of the defence:

Section 32A Defence of innocent dissemination – forum administrator and digital platforms

(1) It is a defence to the publication of defamatory matter if the defendant proves that--

(a) defendant published the matter merely in the capacity as a forum administrator, or the matter was published on a digital platform.

(b) the defendant neither knew, nor ought reasonably to have known, that the matter was defamatory, and

(c) the defendant's lack of knowledge was not due to any negligence on the part of the defendant

(2) for the purposes of subjection (1), a person is a 'forum administrator'

(a) if the matter was published on a digital platform

(b) if the person was not the author or originator of the matter, and

(c) did not have any capacity to exercise editorial control over the content of the matter (or over the publication of the matter) before it was first published.

REBUTTABLE PRESUMPTION

As there is only a presumption that forum administrators are innocent pursuant to s 32A, this innocent dissemination defence is not absolute. This can be rebutted if it is demonstrated that the forum administrator is promoting or sharing defamatory material or content. This can be shown where the forum administrator fails to remove such content without delay after being notified of a request from another user for the content to be removed. Ultimately, it would be inappropriate for the forum administrator to promote or share such content and where they do, the onus should fall on both parties.

TERMS AND CONDITIONS PLACE AN ONUS ON THE INDIVIDUAL

The terms of service between user and platform is a legally binding contract, so internet intermediaries should include user liability in those terms. Users may not always be aware that the content they are posting is defamatory, so it is necessary for terms to cover all posts not just those with intention to defame.

CONCLUSION

In conclusion, a standalone section (clause 32A) should be enacted which creates an innocent dissemination defence for digital platforms and forum administrators. This is because the current 'subordinate distributor' definition is not appropriate considering the nature of the internet. However, this proposed defence is not absolute and if platforms actively promote defamatory content, then liability falls on both author and intermediary.

Question 9: Safe harbour Subject to a Complaints Notice Process

(a) Should a defence similar to section 5 of the *Defamation Act* 2013 (UK) be included in the MDPs?

A defence akin to that of section 5 of the Defamation Act 2013 (UK) should be included, in at least some capacity, into the Model Defamation Provisions. The UK legislation in its present form is already drafted in a manner suitable for Australia The defence is justified in that it does not punish operators for communications they did not publish.

(b) If so, should it be available at a preliminary stage in proceedings, where an internet intermediary can establish they have complied with the process?

Allowing an internet intermediary to raise the defence at an early stage in proceedings will promote efficiency for applicants and respondents. This would be similar to the Anti-SLAPP proceedings followed in many American States whereby certain defamation lawsuits are disposed of at the very early stages.

(c) Should a complaints notice process be available when an originator can be identified? For example, to provide for content to be removed where the originator is recalcitrant?

A complaints notice process is recommended. The legislation should enable the intermediary to self-regulate.

(d) If such a defence were introduced, would there still be a need to strengthen the innocent dissemination defence?

If such a defence were introduced, it may not be necessary to strengthen the innocent defamation defence, as the two defences would work alongside each other.

Question 10: Immunity for Internet Intermediaries Unless They Materially Contribute to the Unlawfulness of the Publication

This question attracted a range of well-developed but contrasting responses. Because each perspective makes a helpful contribution, a range of differing responses is provided.

(a) Should a blanket immunity be provided to all digital platforms for third-party content – even if they are notified about it, unless they materially contribute to the publication?

Yes –

Currently in Australia the liability of internet intermediaries for defamatory content posted by users is somewhat unclear. There is no specific legislation or common law principles that definitively clarifies this issue. To bring greater certainty, Australia should introduce immunity for internet intermediaries. However, this immunity should not be absolute. Exceptions should include when they materially contribute to the unlawfulness of the publication or when the intermediary is notified of the content and refuses to take action to moderate or remove.

Fundamentally, internet intermediaries are not the primary publishers of content. Currently the basis on which these third-party internet intermediaries are liable for the actions of individuals users is confusing and sometimes incoherent. The main limiting devices of liability, mainly intention, passivity and knowledge, are ineffective in drawing a clear distinction for circumstances in which intermediaries will not be held liable.[16] An immunity would bring greater clarity whilst still holding intermediaries liable when they have taken responsibility over the content.

[16] Kylie Pappalardo and Nicolas Suzor, 'The Liability of Australian Online Intermediaries' (2018) 40(4) *Sydney Law Review* 469.

An analogous provision is found in the *Broadcasting Services Act 1992* (Cth). It provides security for online advertising platforms whose hosts were unaware of the existence of the related content. However, for this clause to apply, online content providers must prove a lack of understanding or knowledge of the offensive content posted on their platform. For the proposed immunity, internet intermediaries could be expected to meet certain requirements to meet the immunity. These requirements could include where they are notified of the content and requested to delete or moderate content that infringes on an individual's privacy or reputation, action must be taken. Further, intermediaries must promotedata privacy protection and awareness; and act in accordance with the *Defamation Act* and *Privacy Act* provisions.[17]

We would further suggest that where an intermediary is made aware of content that is defamatory and they do not act to moderate the content, they should be considered to hold responsibility for disseminating said content. It is general knowledge that most intermediaries, and specifically social media websites, do have reporting features that allow users to flag inappropriate and defamatory content. The onus is to ensure that all relevant content on their website is within the scope of their own terms and conditions of appropriate content.

We agree with the discussion paper as to what constitutes 'materially contributing' to the unlawfulness of the publication. The paper suggests that requiring defamatory content on the design of the website platform, intentionally eliciting defamatory comments from users and editing of comments or posts changes the meaning to become defamatory – all amount to materially contributing.

In the above-mentioned exceptions, it could be argued that the intermediary becomes a primary publisher because they are no longer just disseminating user

[17] Australian Law Reform Commission, *Serious Invasions of Privacy in the Digital Era* (Discussion Paper No 80, March 2014).

created and third party content. Primary publishers are liable for defamatory posts and content and as such, when an intermediary becomes a primary publisher, they should be held liable in the same way.

Yes –

There are two fundamental reasons the blanket immunity should be provided to all digital platforms for third party content. Firstly, it is completely unrealistic for these platforms to be aware of the entirety of the content posted on their platform. There are an enormous number of status updates, photos, and videos being posted and shared every second. For platforms to be aware of all this content is simply unrealistic. Platforms are also arguably one step removed from the initial publication and dissemination of the material and immunity would reflect this lessened responsibility.

But what if they are notified about the potentially defamatory content?

Action would still require an enormous undertaking by platforms. It also leaves the door open to platforms mistakenly permitting content that is later deemed defamatory. Would platforms still be liable even if the content was checked? If yes, platforms will begin to remove large amounts of content that may or may not be defamatory. When platforms start removing content without court orders there is a concern that they will be encroaching heavily on freedom of expression, resulting in a chilling effect.

A less obvious issue with not providing the immunity is the effect on entrepreneurs. Australia will fall behind as tech entrepreneurs in this area will face the constant threat of being prosecuted and financially burdened for third party defamatory materials.

Sources:

https://www.justice.nsw.gov.au/justicepolicy/Documents/review-model-defamation-provisions/discussion-paper-stage-2.pdf

https://www.zdnet.com/article/aussie-attorneys-general-considering-blanket-defamation-immunity-for-digital-platforms/

No –

Blanket immunity should not be provided to digital platforms for third-party content even if they do not materially contribute to the publication. Immunity should be provided to organisations if they remove content in a reasonable time once they have been notified of an unlawful publication. If organisations are notified of the content and take an unreasonable amount of time to remove the material, blanket immunity should not be extended. It is necessary to consider the scale of the organisation when considering if they have breached the 'reasonable' time period for removal of content. Even if organisations do not materially contribute to publications, by keeping unlawful content available on their platform they are contributing to the publication by gaining an audience for the material. By allowing the unlawful publication to remain on their website following notification, the digital platform is knowingly assisting in the publication of unlawful material. Due to this, blanket immunity should not be extended once notification has been provided.

Further, there should be a distinction between liability for the dissemination of unlawful content in relation to civil and/or criminal materials. As we have argued against blanket immunity, instead we suggest that there be levels of liability dependent upon the role and function of the internet intermediary; whether they are '"conduits' (network and access providers), 'hosts' (storage) and 'caches' (those who create temporary caches of material to make for more efficient operation of the network)".[18] As Olivier Sylvain suggested in relation to section 230 of the US

[18] Talat Fatima, 'Liability of online intermediaries: Emerging Trends' (2007) 49(2) *Journal of the Indian Law Institute* 155, 159-160.

CDA legislation, the law should be read as providing immunity from liability to sites "that take good faith efforts to take down objectionable materials" not to create blanket immunity.[19] It is this perspective that we suggest should be considered in an Australian law context.

No -

It is impractical to apply a blanket immunity over every digital platform; a more case-centric approach would be more effective.

Even without 'materially contributing' to content, providing a platform where potentially harmful content may be posted should carry with it certain weight and responsibility. Sites that offer interactive platforms, including ones where individuals can comment, should be held responsible (to some degree) for the content that is posted there. This does become a problem; most internet intermediaries are extremely large and home to billions of articles, comments and posts. However, if a platform is aware of potentially harmful or illegal content on their site, they should have the responsibility to deal with that content appropriately. Whether such a responsibility should find its basis morally or legally, remains to be seen.

This is not to say that intermediaries are the publishers or speakers of third-party content published on their platforms. However, it is important to note that because they are providing a large sharing platform, a responsibility of regulation should be placed upon them.

(b) What threshold or definition could be used to indicate when an intermediary materially contributes to the publication of third-party content?

[19] Paul Blumenthal, 'The One Law That's the Cause of Everything Good and Terrible About the Internet' *HuffPost Politics* (online, 06 August 2018) <https://www.huffpost.com/entry/online-harassment-section-230_n_5b4f5cc1e4b0de86f488df86?p2=>.

Despite the disparity of responses to Part (a), the input from the student groups addressing Part (b) was quite uniform.

Internet intermediaries may not control what information is published on their platform, however they do control the ability to filter and remove content.

A threshold finding its basis in Article 14 of the E-Commerce Directive may, as modified below, be appropriate in the Australia, in that an intermediary would be held to "materially contribute" to a third-party publication if:[20]

a) They have knowledge concerning the illegality or confidentiality of the information,

And

b) Upon receiving knowledge or becoming aware of illegal or confidential third-party information, the internet intermediary does not act expeditiously to disable or remove access to that information.

And

c) The internet intermediary did not act to the same standard that a reasonable internet intermediary would act in the circumstances.

Unlike Article 14, the above threshold presumes that internet intermediaries do not materially contribute to third-party publications, unless it can be proven otherwise in accordance with the above requirements. This would prevent numerous claims from appearing before Australian Courts in which internet intermediaries would carry the burden to prove they did *not* materially contribute to any third-party publications.

[20] *Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market* [2000] OJ L 178/1, art 17.

In Canada, Dr Emily Laidlaw and Hilary Young explored what is to be regarded as 'knowledge' for the purposes of a proposed threshold .[21] They maintained that 'knowledge' may be actual or constructive knowledge, and may also be connected to wilful blindness in not instilling the relevant procedures to detect potentially breaching content.

It may also be appropriate that the reasonableness of an internet intermediary's actions be determined in the context of that intermediary's size, scale, capabilities and resources. This would ensure a stable balance between objectivity standards and subjective considerations, giving rise to what is described as a "hybrid subjective-objective test".[22]

Taking on a more nuts-and-bolts practical approach, liability might be approached from the perspective of the website materially contributing to the offending content. This approach has found some support in the United States. By virtue of being notified and subsequently failing to respond, an intermediary has 'materially contributed'. After having been notified of potentially harmful content, a complicit response from a digital platform could be viewed as 'contributing' to its message. Being notified induces the platform to make a choice - to either ignore the notification or react.

The US Court of Appeals decision in *Fair Housing Council of San Fernando Valley v. Roommates.com LLC*, 521 F 3d 1157 (9th Cir 2008) is illustrative of this approach. In this case, the Fair Housing Council argued that Roommates.com was actively participating in unlawful conduct by requiring users to provide details of their age, gender, sexual orientation and other factors. Users could then be filtered according to these factors, which could lead to them being unlawfully discriminated against by other users, therefore breaching anti-discrimination laws.

[21] Emily Laidlaw and Hilary Young, 'Internet Intermediary Liability in Defamation: Proposals for Statutory Reform' (2017) *Law Commission of Ontario* 1, 52.

[22] Ibid 54.

By actively inducing third parties to express illegal preferences, the Council found the US §230 Communications Decency Act immunity would not apply. The Court said:

> '[R]equiring website owners to refrain from taking affirmative acts that are unlawful does not strike us as an undue burden. These are, after all, businesses that are being held responsible only for their own conduct; there is no vicarious liability for the misconduct of their customers, so long as the design of their services does not require users to do something unlawful, and they don't actively encourage users to do something unlawful.'

When deciding whether an intermediary has materially contributed to the publication of third-party content, a time-based threshold could be applied. Here is a proposed action plan:

A. Once any online content has been flagged as being potentially harmful or illegal, a first response from the internet intermediary should be within 48 hours. This is the first step in a chain of responses. This first step allows the intermediary to notify the complainant and the publisher that the content has been flagged. This response should be accompanied by the temporary removal of the content from the platform, subject to review.
B. The second response is the use of an algorithm to filter out known illegal content and have it permanently removed from the platform. For example, the presence of specific words that would suggest illegal content on the basis of legislation (such as discriminatory statements) would result in immediate removal. Ideally, this algorithm would be able to identify and review all forms of content relevant media, including titles and comments.
C. The third response is human review. This step would help identify content that was not as obviously illegal as that filtered out at response two. This would be performed by a team of different stakeholders and may include people of different occupations such as lawyers, healthcare professionals as

well as other platform users. Taking into consideration the mass number of complaints digital platforms receive, a longer time frame for the platform to evaluate the content is necessary - we suggest a 21-day maximum. When reviewing content, it is important to consider the context in which it has been shared. For example, sharing controversial content may be beneficial for educational purposes, but could easily be taken out of context and used for harm. One potential issue with this response is that of differing jurisdictions; what is deemed legal in one country may be contrary to the laws of another, and policing these differences becomes difficult on an online platform.

D. Upon notification of human review, the publishers of the content should be notified. This may take two forms; one result may read 'content has been reviewed and there was an error in the notification', but the other may read 'content has been reviewed and has been deemed unlawful'.

E. If necessary, a fourth response may be introduced; review of the third response by a separate appellate board. If this step is necessary, the decision of the appellate board must be given full authority; any decision made by the board would be final and binding.

(c) If a blanket immunity is given as described above, are there any additional or novel ways to attract responsibility from internet intermediaries?

If a blanket immunity is provided, there could be numerous ways to incentivise internet intermediaries to act responsibly. One such method could include the removal of advertisements on their digital platforms in the case they fail to adequately review and moderate any potentially harmful content posted there. This economic consequence may be the most effective, as internet intermediaries rely heavily on advertising for their funding. Another way to attract responsibility may be to hold them publicly accountable if they fail to deal with potentially harmful content in an appropriate manner. If the intermediary is

publicly shown to be complicit with the posting of harmful and potentially illegal content, this would impact shareholders and individuals thinking of using the platform. This would encourage internet intermediaries to comply with any responsible guidelines, even though they are covered by a blanket immunity.

There are a number of practical and definitive steps that can be imposed on the intermediaries. Internet Intermediaries must become more aggressive at self-regulation. This responsibility can be promoted through:

1. Mitigating brand damage should their platform become a breeding ground for vile content. Internet intermediaries have a branding and economic incentive to be a safe space. Failing this undermines consumer trust and will deter advertisers from using their platform to reach audiences.
2. Being Good Samaritans. Emerging from s230 of the US Communications Decency Act, intermediaries have the discretion to remove obscene or offensive content if done in good faith. There are ethical obligations in business to address wrongdoings. Intermediaries have a moral duty to ensure they are safe and respectful places of public discourse.
3. Standardising practises across the industry to create a cooperative and more consumer-friendly environment where rights and obligations of both parties are clear. An industry coalition on content moderation incentivises the costs of self-regulation by removing considerations of competitiveness in the marketplace.

Additional novel approaches focuses on specific activities. These include:

- Sponsored posts:

Platforms such as Facebook and Instagram offer sponsored posts to businesses/organisations/people that pay to deliver their product or message to a larger audience than their following. They run as short-term campaigns usually no longer than seven days.

By sponsoring a post, internet intermediaries such as Google, Facebook and Instagram ensure that the sponsored posts reach a larger audience than the original post would attract. They do this by placing the target post into the user's newsfeeds. This ultimately means that users will see the post who would not have seen it if it had not been sponsored. Sponsored posts will often appear multiple times within a newsfeed or if a user is watching a video, the clip will pause and the sponsored ad or video will play.

The potential harm with sponsored advertising if left unchecked is that businesses or people could pay Facebook to spread defamatory content or harmful messages. By exposing users to posts that they would not have usually seen without sponsorship, Internet Intermediaries have a direct role in the circulation and dissemination of the content. Furthermore, if Facebook was to be notified of such content yet continue to distribute it, this may be a way of materially contributing, and could attract liability.

Source:

https://fitsmallbusiness.com/facebook-sponsored-posts/

https://techcrunch.com/2018/04/23/facebook-hit-with-defamation-lawsuit-over-fake-ads/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAANaNSBW1Axb2_732Juu6E62A7XtZAAvJ61qeDor7oLuD08X-yBux_orypecWdFGh8mM_5273L4qfg3gWJUZ77mpOW-RC8zNPWRo1GCvoS4hA5h5yZ0Z5K0idcksAyH9Z9wORsxXBu1SbpXVTwbKBzHSio_VwWNceaI07bB1_9Xcx

- Algorithms:

Algorithms used by different platforms are increasingly influencing the way we perceive and experience the world around us. The algorithms dictate everything that users are exposed to on internet platforms and have been criticised for disseminating misinformation and content that exacerbates polarisation. The way

such platforms use and monitor their algorithms could be a potential way to attract liability.

This was highlighted in 2016 with Facebook's 'Trending topics' disaster, where an algorithm was used to promote popular topics. Fake news of the firing of a Fox news journalist was left online for over eight hours. Facebook never found the source of this article and they still use the same algorithm with only a few human editors. This example highlights how a platform's algorithm could attract potential liability for material contribution.

The difficulty with attracting liability for posts shared by automatic algorithms is that internet intermediaries can claim that they had no knowledge of the individual circumstances or posts. If they were to become legally liable for harmful or defamatory content, the platforms would have to redesign their products and re-design algorithms.

Source:

Seth C. Lewis, Amy Kristin Sanders, and Casey Carmody 'Libel by Algorithm? Automated Journalism and the Threat of Legal Liability' (2019) 96(1) *Journalism and Mass Communication Quarterly* 60, 61.

Question 17: Other Issues Regarding Liability of Internet Intermediaries Application of Defences

Regarding proposed defamation defences offered to internet intermediaries, it is unclear whether defences are mutually exclusive or can be employed in combination with one another. To the extent that defences are mutually exclusive, this represents a significant constraint on the ability of internet intermediaries to defend themselves against defamation actions.[23] Internet intermediaries would be forced to stand behind a single line of legal argument, even where other legal arguments may be just as compelling.[24] We believe that the proposed defences could and should be made available as combined defences, forming multiple layers of legal defence for internet intermediaries. This is particularly significant in situations where defamatory material is created by third parties.[25]

The Attorneys-General Review of Model Defamation Provisions Discussion Paper highlights four possible reforms to immunities and defences available for internet intermediaries.[26] Whilst each of these four reforms may be sufficient as independent defences for intermediaries,[27] they would collectively benefit from an ability to be used in conjunction with one another. Essentially, the formation of tiered defences would grant internet intermediaries sufficient protection from liability for content they do not necessarily possess the ability to proactively police. So long as the multiple defences argued are not contradictory, in that proving one defence means the logical impossibility of the other, internet intermediaries should be afforded the opportunity to argue multiple defences. This is likely to take the

---

[23] Australian Government, *Regulating in the Digital Age: Government Response and Implementation Roadmap for the Digital Platforms Inquiry* (Final Report, 12 December 2019) https://treasury.gov.au/publication/p2019-41708.
[24] Ibid.
[25] NSW Government, 'Attorneys-General Review of Model Defamation Provisions- Stage 2' (Discussion Paper, NSW Department of Justice, February 2019) 42 para 3.66. https://www.justice.nsw.gov.au/justicepolicy/Documents/review-model-defamationprovisions/Final-CAG-Defamation-Discussion-Paper-Feb-2019.pdf.
[26] Ibid para 3.67.
[27] Ibid.

form of an 'innocent dissemination' defence, coupled with a secondary defence such as a 'safe harbour' defence.[28]

Internet intermediaries should be afforded the presumption of being a 'secondary publisher' with regards to defamatory materials published by third parties.[29] This would enable intermediaries such as social media platforms and search engines to be able to rely on a primary defence of 'innocent dissemination'.[30] As a first line of defence, these intermediaries will be immune to defamation actions so long as they have acted 'expeditiously' to suppress or remove defamatory material once they have been put on notice,[31] and have not attempted to promote or otherwise curate said defamatory material.[32] This presumption will be rebuttable to prevent intermediaries from acting in bad faith; removing the possibility of intermediaries passively or actively contributing to the harms suffered by a defamed party and then being shielded from the consequences of such actions.

To the extent that the presumption is rebutted, and a defence of innocent dissemination fails, internet intermediaries should then be able to rely on defences such as a 'safe harbour' defence.[33] A 'safe harbour' defence would mean that even where an internet intermediary has failed to act immediately to remove defamatory content posted by third-parties, they may still be able to avoid liability for defamation by actively contributing to the resolution of a complaints process on behalf of the defamed party.[34] This is to say, an internet intermediary may be able to avoid liability so long as they comply with the requests of a defamed party by removing defamatory material, or by connecting the complainant with the originator of the defamatory material.[35] As internet intermediaries often have

[28] Cf *Fairfax Media Publications Pty Ltd v Voller* [2020] NSWCA 102 at [41] per Basten JA.
[29] NSW Government (n 25) 51 para 3.109.
[30] *Model Defamation Amendment Provisions 2020* (Cth) cl 32. See also David Rolph, 'Publication, Innocent Dissemination and the Internet after Dow Jones v Gutnick' (2010) 33(2) *UNSW Law Journal* 562, 580.
[31] NSW Government (n 25) 53 para 3.118.
[32] Ibid. See also *Defteros v Google LLC* [2020] VSC 219.
[33] *Defamation Act 2013* (UK) s 5.
[34] NSW Government (n 25) 59 paras 3.145-3.146.
[35] Ibid 55 paras 3.127-3.130.

sufficient access to both user details and content deletion tools, they are uniquely placed to be able to connect complainants with the originators of defamatory materials so that the complainant may pursue an action against them directly,[36] or to provide the recourse often sought by complainants of removing defamatory material.[37] By allowing intermediaries the opportunity to identify the originator of the offending material in exchange for a waiver of liability, the 'safe harbour' defence would ensure that those who are most responsible for the causing of reputational harm are the ones specifically targeted for legal action.

Through the implementation of a tiered defence system, internet intermediaries will have sufficient protection from being implicated in the defamatory acts of third-parties, so long as they act in good faith in attempting to correct the harms suffered by defamed parties.

Other Considerations:

a. Special liability to be imposed for third-party prohibited content involving protected persons
    i. These may include:
        1. People with physical or mental handicap;
        2. Indigenous Australians;
        3. Elderly people; and
        4. Transgender and intersex people.
    ii. Increased damages available under statute for content involving the above protected class.
b. Implied authorisation regime.
    i. Where a person or entity publishes the content themselves via one ISP, any republication through another ISP cannot open the second ISP to liability.

[36] Ibid 57 para 3.138.
[37] Australian Competition and Consumer Commission (ACCC), *Digital Platforms Inquiry* (Final Report, 26 July 2019) 63 https://www.accc.gov.au/publications/digital-platforms-inquiry-finalreport.

ii. This does not apply to copyright matters where express authorisation will be required.

c. Consider not imposing liability on ISPs

Australia is not in the forefront of media and defamation law reform and liberalization. Imposition of liability may further dissuade revenue and job producing corporations from establishing themselves in Australia

- End of Document -