HERBERT SMITH FREEHILLS

UNMASKING TROLLS: DEFAMATION REFORM FOR THE DIGITAL AGE

09 December 2021 | Australia

Newly proposed ‘unmasking’ obligations will force social media service providers to disclose contact details of defamatory commenters – but what exactly do they need to do and how can they ensure compliance?

On 28 November 2021, the Australian government announced that it would introduce new legislation to address online harms resulting from defamatory comments on social media, with social media platforms to bear obligations to ‘unmask’ users behind such comments.

Shortly afterwards, on 1 December 2021, the Attorney-General’s Department released an exposure draft of the Social Media (Anti-Trolling) Bill 2021 (the **Draft Bill**) for public consultation, including an explanatory paper and detailed explanatory notes for the Draft Bill (the **Explanatory Materials**).

On the same day, the government also announced a House Select Committee inquiry into online safety on social media platforms (the **Inquiry**), including the scope and impact of online harms, the use and impact of industry safeguards, and transparency and accountability of online platforms.

Industry stakeholders may provide feedback regarding the Bill to the Attorney General’s Department until 21 January 2022, as well as to the Committee (which has not yet been formally established at the time of publication, but we anticipate will be established on the next sitting day of Parliament).

WHY HAS THE DRAFT BILL BEEN INTRODUCED AT THIS POINT IN TIME?

In September 2021, the High Court of Australia handed down a landmark decision in *Fairfax Media Publications Pty Ltd v Voller [2021] HCA 27* (***Voller***) regarding the operation of defamation law in Australia in online contexts (see our note here). The Court held that media companies were "publishers" of third party users' comments made on their Facebook pages. This opened the door to organisations or individuals bearing liability for defamatory third party comments on their public social media pages (regardless of whether the party administering the page is aware of or intends to convey the defamatory comment).

In the explanatory materials to the Draft Bill, the Attorney General's Department stated that the federal government did not consider it appropriate for organisations or individuals to bear liability for defamatory third party comments on their social media pages. The Draft Bill proposes to resolve the question of where liability lies for defamatory comments published on social media following *Voller*.

Importantly, in addition to reframing current defamation law in an online context, the Draft Bill also aims to address the unique nature of harm which can result from defamatory comments on social media, where such comments can be made anonymously and amplified.

Under the proposed 'unmasking' obligations, social media service providers will be required to provide contact details (name, email address, phone number and any other specified details) for end users who have made defamatory comments in Australia.

WHO IS AFFECTED BY THE DRAFT BILL, AND WHAT ARE THE KEY CHANGES PROPOSED?

The Draft Bill sets out a number of changes to the current operation of defamation in the context of social media. Most notably:

- **Providers of social media services**, rather than the **social media page owner** (i.e. the organisation or person that manages or administers a social media page) would be publishers for the purpose of defamation – and accordingly, bear liability for third party defamatory comments published on their platforms. This is a clear departure from the position under *Voller*. However, the Draft Bill would also provide a new defence to social media service providers for liability in defamation proceedings, if they have instituted and acted in accordance with the requirements of a prescribed complaints scheme (more on this important defence below).

- **Victims of defamatory comments** would have two new methods of seeking to 'unmask' anonymous commenters, in order to bring defamation proceedings against them: either using a social media service provider's complaint mechanism, or seeking an 'end-user information disclosure order' from the courts. This will provide victims of anonymous defamatory comments with alternatives to the current approach to unmasking the anonymous user who posted the comment, which typically involves an

application to a court for preliminary discovery. In a recent example of that regime, the Federal Court ordered Google to disclose, by way of preliminary discovery, information in its possession about the identity of poster of an allegedly defamatory Google review of a dental practice.

- **End users making defamatory comments** (**commenters**) on social media may be requested to, upon a complaint to the social media service provider, provide consent to the deletion of their comment or the disclosure of their relevant contact details. Their consent will not be necessary for the social media service provider to use the provider's geolocation technology to disclose whether the end user made the defamatory comment from in or outside of Australia.

WHO IS A PROVIDER OF SOCIAL MEDIA SERVICES?

The Draft Bill defines 'social media service' by reference to the definition in the *Online Safety Act 2021* (the **Online Safety Act**), being an electronic service, the sole or primary purpose of which is to enable online social interaction between two or more end users, connects end users with other end users, allows end users to post material on the service and any other specified conditions.

There are a few exclusions provided in the Draft Bill, which clarifies that providers of social media services do not include individuals that merely provide carriage services, or billing or fee collection services in relation to social media.

HOW CAN SOCIAL MEDIA PROVIDERS AVAIL THEMSELVES OF THE NEW DEFENCE TO DEFAMATION LIABILITY?

To avoid penalties and rely on the Draft Bill's new defence for liability in defamation as a "publisher" of comments on their platforms, social media service providers will need to:

- ensure they have a complaints scheme that meets the prescribed requirements set out in the Draft Bill; and
- be able to prove that the complaints scheme was followed in respect of any complaint that has resulted in a defamation proceeding.

In summary, under the scheme, the social media service provider would have to provide the commenter's country location data (being either in Australia or outside of Australia), and where the comment was made in Australia, by seeking to obtain the commenter's relevant contact details and provide them to the complainant. Further detail is in the table below.

Location where defamatory comment was made	Features of complaints scheme offered by social media service provider
In Australia	The social media service provider: • must, within 72 hours of the complaint being made: ◦ inform the commenter of the complaint regarding their comment; ◦ notify the complainant that the commenter has been informed as such; and ◦ disclose to the complainant the country location data of the commenter (being that the comment was made in Australia). • may seek the commenter's consent to remove the comment from the page, and effect such removal. • must, within 72 hours after the outcome of the complaint, notify the complainant of the outcome. • must, where the complainant is dissatisfied with how the complaint was handled: ◦ within 72 hours after the complainant's request, seek the commenter's consent to disclose their relevant contact details; and ◦ if consent is provided, disclose the commenter's relevant contact details to the complainant. Despite the above rules, if the social media provider reasonably believes that the complaint or request does not genuinely relate to the potential institution of defamation proceeding, the social media provider does not need to take any action in response to the complaint or request. Given that, practically speaking, any scheme developed by social media providers will also need to provide guidance for complaints handlers in making (and keeping careful records of) that assessment.
Outside of Australia	The social media service provider must, within 72 hours of the complaint being made, disclose to the complainant the country location data of the commenter (being a statement to the effect that the comment was made from outside of Australia).

WHEN AND HOW WOULD SOCIAL MEDIA SERVICE PROVIDERS BE REQUIRED TO UNMASK A COMMENTER IN RESPONSE TO AN END USER INFORMATION DISCLOSURE ORDER?

Under the Draft Bill, a court would be able to issue an end user information disclosure order to a social media service provider where the applicant reasonably believes (and the court agrees) that they have the right to obtain relief regarding a third party comment in Australian defamation proceedings, and where certain conditions regarding the information available in relation to the comment are met as described below.

When a prospective applicant...	A court may order the social media service provider (or its nominated entity) to provide, within the period specified in the order...
Is unable to ascertain the commenter's relevant contact details	The country location data and, if the comment was made in Australia, relevant contact details of the commenter which the provider (or its nominated entity) holds or has access to
Is unable to ascertain whether the comment was made in Australia	The country location data of the commenter which the provider (or its nominated entity) holds or has access to
Reasonably believes that the comment was made in Australia	

The court may also refuse to make an end user information disclosure order where they are satisfied that disclosure is likely to present a risk to the safety of the commenter.

I AM A SOCIAL MEDIA SERVICE PROVIDER – WILL I NEED TO HAVE A NOMINATED ENTITY?

Only social media services providers that are foreign body corporates would be required to have a nominated entity in Australia, if they have at least 250,000 account holders in Australia or are a specified service under the legislative rules. A foreign social media service provider that should, but fails to, have a nominated entity, may be subject to a civil penalty of up to $550,000.

That nominated entity must have access to the country location data and relevant contact details for comments made in Australia, as well as authority to receive complaints and requests made under any prescribed complaint scheme operated by the social media service.

WHAT OTHER NEW FEATURES ARE INCLUDED IN THE DRAFT BILL?

The Draft Bill includes a mechanism enabling the Attorney-General to intervene in proceedings on behalf of the Commonwealth. This includes in both defamation proceedings and proceedings relating to other provisions in the Draft Bill, such as over an end-user information disclosure order. Commentary by Prime Minister Scott Morrison at the time of announcing the Draft Bill indicated that the Attorney-General would seek test cases and intervene to make clear to the courts how it thinks the legislation should be applied.

The Attorney-General would also be able to make legislative rules with additional requirements for relevant contact details of a commenter and the scope of social media services to which the Draft Bill would apply.

IS THE BILL RELATED TO THE UPCOMING INQUIRY BY THE HOUSE SELECT COMMITTEE ON SOCIAL MEDIA AND ONLINE SAFETY?

During the Inquiry, the Select Committee will scrutinise the Draft Bill, as well as the broader issue of how to deal with anonymous defamatory comments in a digital world. In keeping with the Inquiry’s broad-ranging terms of reference, we expect the Committee also to examine related topics in the online harms and online safety space, including the mental health and wellbeing impacts of social media, the scope and effectiveness of online safety measures adopted by social media platforms, industry transparency and accountability, and other matters. The Government announced that the Select Committee will begin hearings in December 2021 and publish its report by 15 February 2022. As yet, further details have not been published.

Stakeholders can provide feedback on the Draft Bill directly to the Attorney-General by 21 January 2022, but are also encouraged to do so through the inquiry.

HOW DOES THE DRAFT BILL DIFFER FROM THE *ONLINE SAFETY ACT 2021* (CTH)?

As the Explanatory Materials make clear, the introduction of the Draft Bill is part of a broader policy objective of keeping Australians safe from online harm, which includes harmful defamatory content. The centrepiece legislation for online harms is the Online Safety Act – though there are similarities between this legislation and the Draft Bill (particularly the power of the eSafety Commissioner to require tech companies to unmask end-users), under the Online Safety Act, online harmful content and conduct is defined by reference to harmful subject matter or intent, and is unlikely to include defamatory content.

Furthermore, defamation has its own body of law and thresholds, which operate distinctly from the classification and handling of online harmful content under the Online Safety Act; the Act is not intended to provide a means to determine the truth or falsity of statements, but to remove harmful content. Nonetheless, there may be occasions when repeated harmful defamatory comments could be considered cyber abuse, which would be captured under the Act.

WILL THERE BE FURTHER DEFAMATION REFORM IN THE ONLINE CONTEXT?

It is very likely we will see further defamation reform in the near future. Currently, States and Territories, through the Meeting of Attorneys-General, are considering more comprehensive reform of defamation laws to ensure these remain fit for the digital age. In the Explanatory Materials, the government also asserts that this Draft Bill—as a "narrow, targeted packaged aimed at a particular type of harm"—will complement any future reform.

2021 has been a significant year for regulatory reform in the Australian online landscape, with social media platforms faced with new requirements across online safety, online privacy and defamation law. As the government seeks to modernise laws and establish new standards, social media platforms and other industry stakeholders should ensure that they are engaging with emerging proposals in order to create regulation that is fit for purpose in the digital age.

KEY CONTACTS

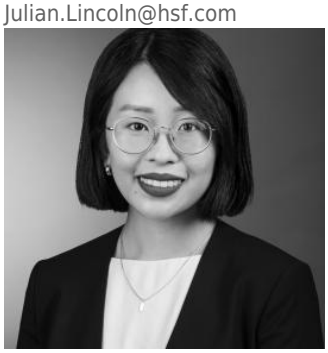
If you have any questions, or would like to know how this might affect your business, phone, or email these key contacts.

JULIAN LINCOLN
PARTNER, HEAD OF TMT & DIGITAL AUSTRALIA, MELBOURNE
+61 3 9288 1694
Julian.Lincoln@hsf.com

TANIA GRAY
PARTNER, SYDNEY
+61 2 9322 4733
Tania.Gray@hsf.com

CHRISTINE WONG
PARTNER, SYDNEY
+61 2 9225 5475
Christine.Wong@hsf.com

KWOK TANG
PARTNER, SYDNEY
+61 2 9225 5569
Kwok.Tang@hsf.com



AUDREY VONG
SOLICITOR, MELBOURNE
+61 3 9288 1387
audrey.vong@hsf.com

TOMAS KEMMERY
SOLICITOR , SYDNEY
+61 2 9225 4939
tomas.kemmery@hsf.com

LEGAL NOTICE