Trustpilot response
Digital Platforms:
Government consultation on ACCC's regulatory reform recommendations

<u>This document sets out Trustpilot's response to the Consultation Paper published in December 2022</u>

Trustpilot A/S ("Trustpilot") welcomes the opportunity to provide input on the ACCC's regulatory reform recommendations and analysis.

We acknowledge the need to ensure a regulatory framework that allows consumers and businesses to continue to be at the forefront in the adoption and innovation of technology in Australia, and to reap the significant benefits of digital development. The online world has the ability to bring consumers and businesses together for positive growth. Ensuring that the right tools and information are being used, and with adequate transparency and safeguards in place to help instil trust in interactions is very much aligned with Trustpilot's mission of enhancing trust online.

A. Trustpilot's online review services

Trustpilot is an online consumer reviews platform headquartered in Denmark, but with a global presence and offices in nine locations, including Melbourne. We bring businesses and consumers together to foster trust and inspire collaboration, and our vision is to be a universal symbol of trust. We are free to use, open to everybody and built on transparency.

Genuine, honest and real experiences shared online are invaluable, both to the people who write and read them, and to the businesses who can use them to understand their customers and improve their offerings. As consumers increasingly purchase goods and services online and read online reviews in that process, we believe it is equally important that they are presented with genuine and honest feedback about legitimate experiences and that they can trust the online review information they see.

On Trustpilot, the majority of businesses and consumers in our community use our platform constructively, and in the way it was intended. For example, the vast majority of reviews (over 94% in 2021) submitted to Trustpilot are *not* flagged by our systems (including our automated fraud detection systems) or our community for breaching our rules, and are displayed on the platform and remain visible. While there is a minority who seek to use the platform in a way that doesn't meet our guidelines and the form and technique of this misuse is continuously evolving, we are, in turn, constantly adapting to meet emerging challenges. We outline some of our proactive work to prevent harm to consumers within the context of the questions below.

Misleading and deceptive reviews can negatively impact consumer choices. We therefore welcome the aim to examine and address the problem of fake reviews. This very much aligns with Trustpilot's own approach as we work to keep our platform a safe and trusted space for the consumers and businesses who use it. We take a strong stance against fake reviews, as we believe that ensuring the integrity of both the platform and the millions of reviews being shared every month is more crucial than ever.

Overall, we agree with assertions in the *Digital platform services inquiry, Interim report No. 5[1]* ("*Interim report*") that *"...trust and confidence in the digital economy is essential to its long-term success"*, and that *"consumers and businesses will only embrace digital opportunities if they are confident that they can trust the technologies and the entities they interact with online."* However, it is precisely this need to ensure trust and confidence that underscores why it is so important to ensure that the problems concerning fake reviews and the potential solutions to address them are well-considered and clearly articulated, and that these set expectations accordingly.

It is also important to take a proportionate approach that bases conclusions about the scale of the problem on a sound body of evidence, and balances negative discourse with fair recognition of the positives and benefits of online reviews, and the power of consumer and business education. As we outline in this response, the prevalence of fake reviews and how they cause harm is not always easily documented or measured, and much of the evidence focuses on online marketplaces or social media. In some cases, the incidence of fake reviews may be overstated. Further, developing potential solutions requires a nuanced and careful approach that ensures that the many benefits of genuine online reviews can also be retained.

B. Trustpilot responses to selected consultation questions

To maximise the usefulness of our contribution to the consultation, we have limited the scope of our response to questions addressing fake reviews and online review services, or digital platforms within the context of fake reviews. This encompasses questions 7-12. Given our global reach across multiple jurisdictions, we have also commented on questions 25-27 of the consultation. We outline our answers to these questions below.

7. Do you agree with the evidence presented by the ACCC regarding the prevalence and nature of harms to consumers resulting from the conduct of digital platforms?

Our response to this question is limited to fake reviews. As we detail below, fake reviews are a legitimate issue that must be proactively addressed to ensure that consumers are

---

[1] ACCC Digital platform services inquiry, Interim report No. 5 - Regulatory reform, 11 November 2022, at page 72.

protected online. However, evidence presented in the *Interim report* on the impact and prevalence of fake reviews could benefit from a more differentiated approach.

Parts of the *report* reference "scams, harmful apps and fake reviews" together, but these are inherently different problems, even if aspects overlap. For example, while there *may* be some overlap, the issues associated with harmful apps are different to those surrounding fake reviews. The platforms these problems may proliferate on, and the underlying root causes for scams, versus harmful apps or fake reviews can be different. It is therefore important to distinguish between these, and take account of their differences. If fake reviews are to be specifically addressed by targeted measures, the characteristics of fake reviews should be considered, as distinct from other, different types of harms. Further, we observe that where specific evidence of fake reviews is presented in the *Interim report*, a majority of the cited texts concern social media, large search engines or online marketplaces. We suggest that in light of this, it may be better to target requirements specifically to these platforms, rather than extending them to *all* digital platforms, without a more detailed analysis of the issues involved.

The impact of fake reviews

It is possible that the impact of fake reviews may be overestimated in some cases. The *Interim report* states:[2] *"By one estimate, 4% of all online reviews are fake, impacting $900 million of spending in 2021 in Australia alone."* This cites an August 2021 article[3] authored by a representative of CHEQ (a business offering Go-to-Market Security solutions and with an interest in highlighting fake reviews left by bots), that states: *"Using official figures and self-reporting by the world's leading e-commerce sites (including Trip Advisor, Yelp, TrustPilot [sic] and Amazon) on average we find that 4% of all online reviews are fake. Translating this into economic impact, the direct influence of fake online reviews on global online spending is $152 billion."* Trustpilot's 2021 Transparency Report does self-report the number of fake *service* reviews removed in 2020 as 2.2M, compared with 40M service reviews submitted in 2020 and this is consistent with a ratio of 5.4%. However, it is an oversimplification to correlate the number of reviews removed as fake with the "direct influence" on consumer spending, as we will go on to explain.

There is also a difference in nature between *product reviews* written about products (such as on Amazon) and *service reviews written about the merchant or trader* (such as the majority of reviews on Trustpilot), and there is not a great deal of evidence detailing how these different types of reviews impact consumer purchases, especially for service reviews.

---

[2] At page 46.

[3] Cited as "World Economic Forum, Fake online reviews cost $152 billion a year. Here's how e-commerce sites can stop them, 10 August 2021. Estimates based on self-reporting by large e-commerce sites including Trip Advisor, Yelp, TrustPilot and Amazon and other sources."

Compounding this, not all platforms define fake reviews uniformly.[4] Definitions can vary between platforms, and this is logical where the type of reviews collected and mode of collection differ. There can also be different incentives for posting fake reviews which differ for online marketplaces versus independent online review platforms; on marketplaces, suspension of a retailer for posting fake reviews can have a *direct* impact on their ability to sell via the platform, and therefore unequivocally affects competition and market share.

Finally, on Trustpilot, we note that a majority of the fake reviews counted in our 2021 Transparency Report are *removed* by our automated systems, while others are removed manually. This means that many reviews will have been online for a limited time. Further, in April 2022 we introduced a short delay of up to two hours from when a review is submitted to when it is published onto the platform. This delay allows us to run our automated technology in a different and staggered timeframe, and provides an even greater opportunity to prevent fake content from ever being seen. Only once our automated technology has analysed the review, will it appear on the platform, becoming visible to everyone. This safeguard further complicates the ability to quantify the impact of fake reviews, since some fake reviews may be removed before they are ever seen by consumers.[5] Moreover, it challenges the assertion made by CHEQ in relation to the impact of fake reviews on global online spending.

The prevalence of fake reviews
The mode of review collection is relevant to people's perceptions about the validity of reviews, and the fake reviews they might encounter.

"Open" review platforms such as Trustpilot collect reviews from reviewers that may not always be recognisable to businesses, which *can* lead businesses to conclude that these reviews are authored by people who have not had an experience with them. While fake reviews written by competitors is a legitimate concern, on Trustpilot we do find that some businesses can be quick to conclude that a negative review (a 1-star or 2-star review) must be written by a competitor. This arises especially in cases where businesses don't recognise the reviewer's username (which can be pseudonymised by the reviewer to protect their privacy) or where the reviewer may not have made a purchase but has had a genuine experience.[6] However, while we do see allegations of suspected

---

[4] For example, Trustpilot's relatively broad definition of "fake reviews" includes not only a review that is not based on a genuine experience or is not about the business being reviewed, but also a business leaving a review on its own Trustpilot profile; a review that has been paid for in an effort to manipulate a business's rating on Trustpilot; a review left on a competitor's Trustpilot page in a deliberate attempt to undermine their rating on Trustpilot; advertising or promotional messages of any kind that are disguised as reviews; or a review that has been removed from the platform because it contains harmful or illegal content.

[5] We note that on Trustpilot, reviewers are notified when their reviews are removed prior to publication and given the opportunity to challenge any decisions they believe are incorrect.

[6] For organic service reviews on Trustpilot, a lack of purchase does not necessarily mean the person cannot write a review, as long as they have had a genuine service experience with the business.

competitor-penned reviews, it is rare that—on exploration—these claims can be fully substantiated with clear evidence.

In light of this, it is possible that the incidence of "fake" reviews within this category is overestimated. One of the more compelling figures on this point is a survey cited by the ACCC's *DPSI September 2022 Report - Discussion Paper,*[7] finding that 38% of business owners reported an experience of an untrue report posted on their listings and 33% reported a competitor had left a review on their listings. A plain reading could imply that these figures are self-reported, but unfortunately, the source document, a 2017 *Digital Reputations Report* by True Local, is no longer downloadable.[8]

Even if a reviewer is not recognisable to the business, they may in fact have had a valid service experience,[9] (and they may also be willing to verify their identity using our optional consumer verification tool and/or provide documentation of their experience if requested.) In 2021, we reported that the majority of reviews flagged by businesses were for the reason "not based on a genuine experience." Over 431,000 reviews were flagged globally for this reason. The majority of these flagged reviews were 1-star or 2-star reviews, but not all of these were deemed as *valid* flags by our Content Integrity Team.

Consumers may also perceive the number of "fake" reviews to be higher than it is. We provide an easy-to-use mechanism for consumers to flag reviews to us that they believe are not based on a genuine experience. Our data shows that globally, this flagging option is used relatively frequently, but consumers or readers of reviews may report reviews that they speculate "look fake", based on perhaps the length of the review, the language or words used, or the tone, and this is not always reliable. As explained in our blog post from 2020, "How to spot a fake review: there's more to it than you think" these indicators do not necessarily allow people to accurately identify fake reviews, and tend to be less effective than analysis of behavioural patterns using the hundreds of data points behind every review. (We note this is also consistent with statements made in a document cited in the *Interim report—"CPRC, Online reviews: a guide not a gospel, December 2019"*, which asserts that the majority of consumers are overconfident in their own ability to distinguish real from fake reviews.)

In 2021, consumers globally flagged over 39,000 reviews to us for "not based on a genuine experience," and these reviews were assessed by our Content Integrity Team. Not all of these submissions were deemed valid, and in fact, across all our flagging reasons,

---

These service reviews can be valuable to businesses, as they can assist with helping to understand why consumers choose *not* to complete purchases with them.

[7] ACCC. February 2022. Digital Platform Services Inquiry, Discussion Paper for Interim Report No. 5: Updating competition and consumer law for digital platform services, at page 52.

[8] The link to download the report at https://www.truelocal.com.au/blog/report generates a 404 error.

[9] We include in our *Guidelines for Businesses* for service reviews: "*Don't recognize a reviewer? We give reviewers the freedom to choose their own username, so you might not always recognize their name as a customer of yours. But that doesn't mean they haven't had a genuine, recent experience with your company…*"

consumers only had a 16% flagging accuracy rate. Further on this point, page 77 of the *Interim report* states: *"A survey of 1,000 Australian consumers commissioned by Reviews.org reported that 52% of respondents believed they had fallen for fake reviews and 28% did not trust online reviews"*, citing *G Dixon, 'More than 50% of Australians Believe They've Fallen for Fake Reviews', Reviews.org, 13 August 2021*. However, the source article also states that: *"26% of Aussies did not correctly identify a fake review,"* which is consistent with our observation that fake reviews can be difficult for consumers to detect based on review content alone. This also calls into question the accuracy of self-reporting by respondents claiming they have fallen for fake reviews.[10]

Alongside using sophisticated technology that can assist in identifying patterns of fake reviews, Trustpilot's Content Integrity agents who assess reviews must also carefully balance competing rights, such as the right to carry on a business with the right to freedom of expression. Businesses can indeed be negatively impacted by reviews and should flag reviews where they have evidence that these are maliciously posted by their competitors (and on Trustpilot, where this is found, such reviews would be removed as fake reviews in breach of our guidelines). However, it is also important to balance a *suspicion* of fake reviews against the negative effects of silencing potentially legitimate consumer voices. This is a delicate balance, and one that we continue to refine via our policies and processes, as well as being as transparent as possible about these without providing details that will allow bad actors to game our systems.

Given the evidence demonstrates that both consumers and businesses struggle to accurately identify fake reviews, it is vital that any future policy recognises and takes account of this issue, both with regard to the policy route chosen and to the data it is founded on.

8. Do you agree with the ACCC recommendation to introduce targeted measures on digital platforms to prevent and remove scams, harmful apps and fake reviews? Are there any other harms that should be covered by targeted consumer measures, for example, consumer harms related to the online ticket reselling market for live events?

Protecting consumers from fake reviews is very much aligned with Trustpilot's mission of enhancing trust online. In principle, we agree with some of the recommendations, especially where platforms are provided with the flexibility to implement these in a way that is most appropriate to the type of service they provide and the specific functionalities involved. However, certain aspects of the recommended requirements should be limited to specific platforms.

---

[10] Reviews.org does not explain further how this was determined for their study - there is little detail regarding methodology as the article states: *"We surveyed 1,000 Australians about online shopping reviews and fake reviews, then analysed the results for this report"*. We note also that the article quotes an unnamed "spokesperson for the ACCC" about the prevalence of reviews that are not legitimate, which creates a circular reference when cited in an ACCC report.

Trustpilot

We note that the measures recommended by the ACCC coincide with those that Trustpilot already undertakes. For example, these include:

- **Notice-and-action mechanism to report fakes:** This is a useful tool for platforms, but given the diversity of models, platforms should be allowed flexibility in exactly how such mechanisms are implemented and operate.

  On Trustpilot, we are continually refining our notice-and-action mechanism to enhance its effectiveness, as part of our three-pronged combination of Technology, Community and People to tackle fake reviews. This is overseen by a Trust & Transparency Team made up of over 90 employees including agents, investigators, lawyers, technology experts and communications and training specialists. We provide accessible avenues for our community to help us promote and protect trust on the platform by reporting fake reviews, where our notice-and-action mechanism allows consumers and businesses to flag a review to us at any time if they believe it breaches our guidelines or is fake. We also provide a "whistleblower" functionality, for people to confidentially report any problems to us, and this can include suspicions about fake or biassed reviews. All reported reviews are considered by our team of experienced Content Integrity specialists. Supporting our Content Integrity Teams are a Fraud and Investigations Team, who consider and act on any particularly complex situations. Both teams have access to powerful tools to help them examine review patterns for further anomalies.

- **Review verification disclosures:** Transparency is a core principle underpinning how we operate. We believe that platforms should provide information about how reviews are collected so that consumers can decide how much weight to afford to them.

  We strive to provide as much useful information as possible to help consumers and businesses understand our platform (including how it works and our business model), as well as our processes and policies for preventing, detecting and removing fake reviews. This includes devoting considerable resources to understanding the consumer journey and user experience in order to provide useful, bite-sized information where it is most relevant. This encompasses (but isn't limited to) information about the platform and review collection process on each company profile page, as well as review labels with hover-over explainer texts, and Company Activity pages that break down the source, volume and timing of reviews accumulated by each business. We link to Help Center articles explaining the different types of reviews and how they are collected and why some are marked "verified" but others are not, measures taken to protect the integrity of the content on our platform, how our fraud detection systems work, etc.

- **Public reporting on mitigation efforts:** We take the view that transparency reporting is a useful tool in helping consumers and businesses understand the extent of fake reviews, and the effectiveness of the measures being taken to combat it. However, platforms need to be afforded some level of flexibility regarding the level of detail they provide, so as not to undermine efforts to protect the integrity of their content.

  On Trustpilot, we take a proactive approach to transparency and report annually on our mitigation efforts via a detailed Transparency Report. Our third such report is due to be released in Q2 of 2023. Our report from 2022 details how we operate and safeguard our platform, how we combat fake reviews and misuse of our platform and how we're tackling key issues raised by consumers. We break down how many fake reviews were detected and removed, against overall figures about how many reviews were written and submitted to our platform during the year, whether reviews were collected by manual or automated means, how many reviews were flagged by consumers or businesses, with a breakdown of which reasons, how many reviews were marked as "verified" etc. We also detail efforts to protect consumers by displaying public warnings where a business breaches our guidelines (for example, by procuring fake reviews), and explain how we display alerts where businesses are subject to regulatory attention.[11]

- **Verification of advertisers of financial services and products:** Requirements regarding advertisers of financial services should be limited to platforms who host third-party advertisements.[12]

- **Verification of certain business users:** The recommended protections for "digital platforms" include verification of certain business users, including advertisers, app developers and merchants in order to minimise scams and harmful apps. In this context, the *Interim report* refers to verification requirements mandated for certain digital platforms under the EU Digital Services Act. As underscored in the *Interim*

---

[11] If businesses don't stop behaviour that breaches our rules, we'll place a red, publicly visible "Consumer Warning" on the business's profile page shown on Trustpilot. Consumer Warnings are displayed on a profile page for a fixed period of time, and we will periodically review it to see if a business is continuing to misuse our platform, only removing it until misuse has ceased, and a reasonable amount of time has passed: this helps to ensure that consumers have been made aware of a business's attempts to mislead consumers, or otherwise misuse reviews on Trustpilot. We also explain how we place what we call "Consumer Alerts" on business profile pages when we believe there is information that consumers should know. This includes where businesses are under the attention of a regulatory body, such as the FCA. We add this alert when we become aware of a new investigation or warning issued by a regulatory authority, or when we're otherwise made aware of a company facing regulatory scrutiny. These alerts stay online for as long as is necessary to raise awareness, or until the regulatory body removes its own notification about the company.

[12] We do not show ads on our platform, other than in the U.S. Trustpilot operates what's commonly called a 'freemium' business model, where we provide our basic services to both consumers and businesses for free, and generate income from businesses subscribing to our premium software services (known as a SaaS model).

![Trustpilot logo]

report, these requirements apply **specifically** to online marketplaces. Accordingly, if any Australian requirements on this point are introduced, we suggest they should similarly be limited to online marketplaces. Other digital platforms should be afforded the flexibility to voluntarily adopt similar measures in a way that best correlates with the type of service offered. Allowing flexibility for platforms to tailor consumer safeguards and protections to suit their individual model can result in effective, helpful and appropriate measures.

For example, at Trustpilot, we facilitate verification of both consumers and business users, but in a way commensurate with the nature of our services and our freemium model. We learned that consumers are looking to gain further insight on the businesses they see reviewed, and many are willing to act on behalf of one another, especially when it comes to the challenges of misinformation online: in fact, 85% of people in the UK and the U.S. said they would be willing to prove their identity to protect others from misinformation online. Both the search for greater context, but also this willingness to help one another—a sense of altruism within our community—led us to introduce the ability for both consumers and businesses to voluntarily verify their identities. The process requires both parties to safely and securely share a copy of their government-issued photo ID, as well as take a selfie. We're using the same technology used by banks, healthcare providers and educational institutions. This is optional and allows everyone to play their part in building an even more trusted community on Trustpilot, and we are now able to highlight where businesses and consumers have done this, as a 'trust signal' to those looking at the information. We released this technology globally in the early part of 2022 and quickly had over 49,000 reviewers successfully verify their identity (this number has now increased), enriching our community of reviewers on Trustpilot.

For businesses that have claimed their profile page on Trustpilot and provided extra verification, we display which details they have chosen to verify, and those they have not, on their company profile page hosting reviews. For example, this can be proof of identity, contact details, registered domain name ownership or bank account details.

- **Mandatory internal dispute resolution standards and an independent external ombuds scheme:** See our answer to **question 10**, below. In principle, we support the creation of an internal dispute resolution mechanism (and note that many platforms operating globally will already have such systems in place), but we do not believe that an independent external ombuds scheme is necessary. If an ombuds scheme is implemented, it should be limited to selected digital platforms, and should only address system-level problems, rather than adjudicating individual disputes.

Trustpilot

Other harms that could be covered by targeted measures

We have identified several further aspects which might be highlighted in order to assist digital platforms in tackling fake reviews:

- The *Interim report* cites instances of websites and groups offering fake reviews. We suggest it could be useful to ensure clear prohibitions against 'hosting the offer or advertisement' of fake reviews so as to help secure action from companies and sites who host these offers or ads. This would assist in delivering a more holistic approach to tackling fake reviews.

- While misleading and deceptive conduct is already prohibited under the Australian Consumer Law (ACL), publicising the consequences of clear-cut and deliberate misrepresentation is useful. This includes, for example, businesses blocking negative reviews from being published to their site, while claiming that all reviews are displayed: e.g. the *Fashion Nova* case in the U.S., and the ACCC case against Meriton Apartments for preventing consumers from submitting negative reviews on a review site, by "masking" the emails of customers they suspected would provide negative reviews. Measures highlighted could also extend to funnelling negative reviews away via an app or a model that incorporates a complaints process where a "bad review" is detected just before/at the point of posting, and redirects to an internal feedback form instead.

- Potentially it could also be helpful to clearly prohibit scraping of online review sites, and review hijacking (where reviews about another product/service are copied and used in a different context).

Finally, it should also be noted that in contrast to the **problems** that can be caused by the misuse of reviews, this should be viewed in light of the many **benefits** that reviews can also deliver. Consumers also routinely share anecdotal evidence with us about how reading Trustpilot reviews has helped them avoid transacting with an untrustworthy business where their view is that they would have suffered a loss if they had done so. While it is important to address the problems of fake reviews, this should be done in a way that is proportionate to the extent of the problem, and balanced against recognition of the fact that where safeguards are employed, online reviews can be a useful tool for both consumers and businesses.

8.1 Is the notice and action mechanism proposed by the ACCC for these consumer measures appropriate? Are there any alternative or additional mechanisms that should be considered?

The *Interim report* outlines (at page 83) features that notice and action mechanisms should include.[13] For online review platforms, we note that while prompt action and clear

---

[13] "Notice: platforms must provide user-friendly mechanisms for individuals and entities to report scams, harmful apps, or suspected review manipulation; Action: platforms must promptly respond to notices, for example, by removing suspected […] fake reviews or providing advice about the basis on which the content is permitted; Communication: platforms must promptly notify the reporting person

communication are desirable, this must also be balanced against the evidence about consumers' abilities to accurately detect fake reviews, and the risk of providing such detailed information which could assist bad actors in gaming the system. Where our Content Integrity agents have investigated suspected fake reviews, the relevant parties are usually notified and we strive to be as transparent as possible. But as a practical measure, and to protect the integrity of our platform, it may not be possible to reveal all of the details of each investigation and the workings of the technologies relied upon to make an assessment.

On information sharing, we recognise that given the dynamic pace of change, a comprehensive and holistic approach across the online landscape is potentially useful, and could, for example, include cross-industry cooperation. We are already seeing some positive examples of multi-actor collaboration in other jurisdictions. As one example, in Europe there are forums that facilitate collaboration and coordination between different actors across different sectors (such as insurance, banking and law enforcement, through to online reviews). While respecting relevant privacy obligations, these potentially allow learnings to be shared to uncover and combat emerging practices in fraud and misuse, and employ tactics to target them at different parts of the consumer journey. Such a coordinated approach could be more effective than, or operate as a complementary initiative to, regulatory changes, which tend to be less dynamic against the fast pace of development in tech and also of bad actors, and more targeted to address limited aspects of the broader picture.

We would also raise concerns about redress, other than removal of fake reviews content where this is relevant. The *Interim report* provides that *"platforms should be required to provide redress to users who have been harmed by the platform failing to meet its obligations under these measures,"* but it is unclear what this should involve, and this is likely to differ depending on the circumstances. In some cases, it will be difficult to assess harm caused directly, and caution must be taken to ensure that any requirements are proportionate, and that they build in adequate safeguards.

Education
We also believe that there may be a greater role to explore in educating consumers about how to use reviews that can help lessen the impact of fake reviews, and increasing overall knowledge about how different digital platforms work. For example, we explain in our Help Center article "How can reviews improve your online shopping experiences?" that reviews

---

and potentially affected consumers of processes and actions undertaken in response to the report. [...]; Information sharing: platforms must promptly share information about identified issues with other platforms and relevant agencies to aid consumer protection efforts. This aspect is particularly important given that many scams operate across platforms and are likely to migrate to less protected platforms or services; Redress: platforms should be required to provide redress to users who have been harmed by the platform failing to meet its obligations under these measures (for example, by failing to act within a certain time after being notified of a scam on the platform)."

are one of many helpful tools available to assist consumers, but they should always be used in conjunction with other resources and should never replace common sense.

9. What digital platform services should be captured in the ACCC's recommendation?

The *Interim report* states that *"[s]ome of the harms identified by the ACCC are not limited to conduct by digital platforms with substantial market power. The ACCC has observed a range of harms to businesses and consumers resulting from a broader range of digital platforms: failing to take sufficient steps to prevent scams, harmful apps, and misleading or fake reviews from proliferating on their services…"*.[14] As stated above, we acknowledge that fake reviews are a legitimate concern and must be addressed with adequate safeguards and an appropriate level of transparency. However, it is worth noting that much of the evidence cited throughout concentrates on the largest social media platforms[15] or marketplaces.

We also note that Trustpilot has already implemented safeguards and measures that align well with the ACCC's *Recommendation 2*. To further develop the proposed approach, it is important to recognise the diversity of the platforms in scope and that a one-size-fits-all approach to policy may therefore not be appropriate. We therefore suggest that certain elements outlined in the recommendations should be limited to specific types of platforms. We would also caution against extending specific and targeted measures to *all* digital platforms with regard to fake reviews. A more nuanced approach would be more appropriate.

With any proposals that are made in line with the recommendations, we urge the adoption of a careful and tailored approach so as to effectively target the issues which have been identified. Where possible, opportunities should also be taken to align with the international measures that, over time, are shown to be most effective.

10. Is a new independent external ombuds scheme to resolve consumer disputes with platforms warranted? Can any or all of the functions proposed for the new body be performed by an existing body and, if so, which one would be most appropriate?

We suggest that, in line with the international approaches cited, an ombuds scheme may be more appropriate in specified situations, where it is limited to particular digital platforms. We suggest that it should not extend to *all* types of digital platforms.

In the online reviews space, we take the view that recourse to an independent external ombuds scheme for resolving disputes should not be necessary. This is especially so where

---

[14] At page 44 of the *Interim report*.
[15] For example, at page 79.

platforms proactively employ proper safeguards in their systems and processes, provide an adequate notice-and-action mechanism to report reviews and have a well-functioning internal decision dispute process, and provide transparent information about their mitigation efforts.

On this point, we note that Trustpilot has already implemented an easily accessible internal decision dispute process for review of decisions made by our Content Integrity agents, including where a review has been removed, to contest the outcome of a review that was flagged but not removed, or to dispute the blocking of a user account, among other things. Each ticket is escalated to one of our specialist team members for review, with updates provided along the way and the outcome communicated. In line with our commitment to trust, our team aims to be as transparent as possible, always works in line with our policy and processes, and takes a neutral approach, treating consumers and businesses equally. However, it is not always possible to communicate every detail of an investigation, especially where there are indications of fraud or fake reviews—and this can be complicated by complexity, especially if users are unclear about how the platform operates.

If implemented, an ombuds dispute resolution scheme should be limited to addressing system or process-level issues, rather than taking individual decisions on content. We would caution against the individual review of claims *"which have not been resolved to the consumer or business user's satisfaction"*, without further and careful qualification, since this could risk drawing out individual disputes and, in some cases, diverting considerable resources to unfounded claims.

As stated above, for individual cases, it may not be possible to reveal the full details of an investigation if there is a risk that revealing such information could expose services to bad actors manipulating their systems. A level of discretion may be required with regard to revealing the details of any investigation, and this can, in some cases and despite explanation, be perceived as a lack of transparency rather than a legitimate precaution. Further, for double-sided platforms in between consumers and businesses, there is a greater risk that complaints involving fake reviews will necessarily be resolved to the satisfaction of one party but not the other—either the consumer whose review remains online, or the business who has the review removed. Investigations into fake reviews can be complex and as noted above, assessments as to whether suspected fake reviews should be removed involve balancing competing rights on the part of the business (the right to conduct a business) with that of a consumer (such as freedom of expression). It is necessarily the case that in all instances, one party will be dissatisfied with the outcome.

Further, the heterogeneity of different digital platform models and the specialist technologies employed means that analysis may potentially warrant significant expertise on the part of the ombuds. If they are empowered to make binding decisions, such as determining that certain reviews are fake and must be removed, therefore overturning platform decisions, it could also risk further undermining public trust in platform processes.

The *Interim report* references the EU's Digital Services Act,[16] stating that *"online platforms are required to implement out-of-court dispute settlement mechanisms"*, and *"[o]nline platforms captured by the Digital Services Act must also engage in good faith and are bound by the decisions taken by the external body."* However, it is worth noting that the final text of the adopted Act[17] also extends good faith obligations to complainants: *"Both parties shall engage, in good faith, with the selected certified out-of-court dispute settlement body with a view to resolving the dispute"*, [emphasis added]. It also includes the caveat that, *"online platforms may refuse to engage with such out-of-court dispute settlement body if a dispute has already been resolved concerning the same information and the same grounds of alleged illegality or incompatibility of content,"* and that in fact *"the certified out-of-court dispute settlement body shall not have the power to impose a binding settlement of the dispute on the parties"*, [emphasis added]. Among other things, this addresses concerns raised during the Digital Services Act negotiation process that spurious or unfounded claims could be brought, which would not accord with the purpose of the provisions. In the context of fake reviews, it is also important that disputes are limited to addressing reports that are legitimate, as opposed to those that have already been investigated, or that are unsubstantiated, or even unfounded. We suggest that such safeguards are valuable and necessary components of any new requirements.

11. The ACCC recommends these requirements to apply to all digital platforms, do you support this? If not, which requirements should apply to all platforms, and which should be targeted to certain entities?

In line with the EU's Digital Services Act (recognised in the *Interim report*), requirements for verification of businesses should be limited to online marketplaces.[18] In our view, the European legislation has correctly limited the application of this requirement to online marketplaces, in light of that particular business model.

The requirement should not be extended to *all digital platforms* since it is neither practicable, nor proportionate. As stated above, this would allow other types of platforms flexibility to introduce measures that are most appropriate for the type of services they offer and platform functionalities.

---

[16] At page 100.

[17] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act), Article 21.

[18] At page 84 of the *Interim report*, it states that the Digital Services Act's 'Know Your Business Customer' obligations require *platforms that enable users to make distance contracts with traders,* to receive identity information from certain traders.

Trustpilot

12. If the above processes are introduced, is the Australian Consumer Law the appropriate legislation to be used and what should the penalty for non-compliance be?

Any penalties for non-compliance should involve an escalation process before sanctions are implemented. Platforms that are proactively taking steps to protect consumers and are willing to cooperate with enforcing bodies on improving their processes should not be penalised in the same way as those who are not.

25. Should Australia seek to largely align with an existing or proposed international regime? If so, which is the most appropriate?

Leveraging international regulatory approaches and overseas industry undertakings once they are developed is likely to better align Australia with larger markets, and can also deliver the benefits of learning from the rules implemented in other jurisdictions.

Taking a differentiated approach could risk that digital platforms—and particularly those that are smaller in scale and size as compared to 'big tech'—either have to implement forms of geo-blocking (which may or may not be possible), or potentially opt to not have a presence in the Australian market. This means that Australian consumers in particular could risk losing out on some of the benefits digital platforms can provide.

26. What are the benefits and downsides of Australia acting in advance of other countries or waiting and seeking to align with other jurisdictions?

Given the high volume of regulation that will need to be implemented by digital platforms with global reach in the coming period, seeking to align will have the advantage of positioning Australia as a market where compliance is practicable and more achievable in terms of cost.

27. Are there any particular aspects of the ACCC's proposed regime that would benefit from quick action or specific alignment with other jurisdictions?

Rather than quick action, we suggest there could be advantages with refining the scope of rulemaking to ensure it is specifically targeted to the most relevant types of platforms, and is proportionate to the harms involved.

It is also possible that before moving to rulemaking, there may also be benefits in seeking to maximise the effect of other steps, such as the ongoing education of businesses about their obligations, and educating consumers about deceptive practices, or exploring whether

codes could be useful. As with most technology regulation, investment in new rules may face the challenge of being quite static in a fast-paced environment, ultimately making it an uphill battle to keep pace with developments in the area. We therefore suggest that pursuing a package of targeted measures may be resource-effective in reaching the desired policy goals and could have the benefit of being quite flexible and dynamic with the ability to adapt to an ever-changing area.

As stated above, we appreciate the opportunity to provide comments on the suggested reforms. We would be happy to share further information to inform and support this work, and would also welcome the opportunity to discuss these issues further with the ACCC and government, and participate in any workshops on relevant topics.

Trustpilot
14 February 2023