



**DMRC Submission in response to the exposure draft of the Communications Legislation  
Amendment (Combating Misinformation and Disinformation) Bill 2023**

Prof Axel Bruns, Dr Ariadna Matamoros-Fernandez, Hannah Klose, Louisa Bartolo, A/Prof Stephen Harrington, Ned Watt, Nadia Jude, Prof Daniel Angus, Prof Dominique Greer, Rhyle Simcock, A/Prof Timothy Graham, and Dr Aaron Snoswell.

**About QUT's Digital Media Research Centre**

The Digital Media Research Centre ('DMRC') at the Queensland University of Technology is a leading research centre in digital humanities and social science research with a focus on digital communication, media, and the law. Our research programs investigate digital inclusion and participation, the digital transformation of media industries, the growing role of AI and automation on digital societies and the role of social media in public communication. Scholars at the DMRC are particularly engaged in research that is directly relevant to much of the subject matter raised in this Bill, including ongoing work to investigate misinformation and disinformation on digital platforms and the appropriate approaches to regulate the harms arising from it.

**Overview**

We thank the Department of Infrastructure, Transport Regional Development, Communication and the Arts for the opportunity to make a submission in response to exposure draft of the Communications Legislation Amendment (Combating Misinformation and Disinformation) Bill 2023. We recognise the importance of tackling harmful online misinformation and disinformation on digital platforms in Australia. However, we are concerned with a number of aspects of the Bill – which we outline in this submission – and make the following recommendations in response to those concerns:

- The distinction between 'misinformation' and 'disinformation' provided for in clause 7 of the Bill be removed. Instead, the central criterion should be whether the continued circulation of that content could cause or contribute to serious harm, as defined in clause 2 of the Bill.
- The definition of 'good faith' be tightened to include satire and parody as content that has the potential to harm.

- Carefully examine the definition of a “news organisation” and consider how exclusions for “professional news content” could be exploited by misinformation agents who self-identify as “journalists”.
- That data gathering powers be strengthened to provide more robust and reliable data access to accredited researchers, civil society actors, and journalists who act in critical oversight roles in terms of detection, tracking, and reporting on problematic online content, and platform responses to such content.

### **Overall scoping and problem definition**

While we welcome any attempt to improve the quality of Australia’s public discourse and tackle the various problems and threats posed by misinformation, we believe that the Bill takes an overly narrow view of the problem of misinformation, positioning it as largely a technological problem that occurs online, generated by ‘bad actors’ using new technologies (AI, automated bots, deepfakes) to produce ‘false, misleading or deceptive content’ on digital platform services. The Bill positions these actors (nefarious online users) and this content (false, misleading or deceptive) as directly contributing to (and in some cases ‘causing’) a laundry list of ‘serious harms’. While we strongly support a harm-based approach to the issue, we are concerned at the techno-centric focus to the underlying issue. Whilst ‘bad actors weaponising new technologies’ is a serious problem that warrants attention, emerging research in the field of misinformation and disinformation suggests that this focus is too narrow and potentially misdirected.

**Recommendation 1:** If the purpose of the Bill is to ‘combat misinformation and disinformation’ and its relationship to the ‘serious harms’ you list, a more holistic and locally relevant conception of the problem of misinformation and disinformation is required. For example, this includes foregrounding that some of the most harmful forms of misinformation and disinformation in Australia have:

- Co-opted emergency events, targeted historically marginalised communities, and worked to exacerbate existing inequalities. For example, the #ArsonEmergency and #DanLiedPeopleDied campaigns exploited the 2020 Australian bushfires and coronavirus pandemic;<sup>1</sup>
- Involved deep mistrust in institutions, stemming from diverse socio-historic factors, such as

---

<sup>1</sup> Timothy Graham et al, ‘#IStandWithDan versus #DictatorDan: The Polarised Dynamics of Twitter Discussions about Victoria’s COVID-19 Restrictions’ (2021) 179(1) *Media International Australia* 127.

historic and continued failures in government policy, services, programs and initiatives;<sup>2</sup>

- Existed within and been amplified by government authorised content (The Voice pamphlets<sup>3</sup>, Craig Kelly's electoral signage);
- Existed within and been amplified by influential actors, such as celebrities, politicians and professional news media;<sup>4</sup>
- Involved an interplay between politicians, partisan news media and social media, where online audiences are both co-opted and aggravated by sensationalised and misleading headlines, contributing to polarisation and a systemic breakdown of the Australian media ecology;<sup>5</sup>
- Existed in (and been excused as) entertainment, satire or parody;<sup>6</sup>
- Not actually involved explicitly false information, but weaponised the context of genuine information in a manner that misleads audiences, regardless of intentionality. For example, the misquoting of arson statistics in leading Australian newspapers to frame climate change as inconsequential for bushfire disasters.<sup>7</sup>

A definition of the problem that includes these factors would change the form and scope of content exemptions within the Bill. It would also lead to a broader range of proposed solutions that reach beyond incentivising platform action designed to circumvent the way misinformation and disinformation 'travels' on certain digital platforms. For example, multi-stakeholder approaches that seek to build stronger and more resilient community structures and civic participation.

---

<sup>2</sup> Bronwyn Fredericks et al, 'Burden of the Beast : Countering Conspiracies and Misinformation within Indigenous Communities in Australia' (2022) 25(1) *M/C Journal* <<https://espace.library.uq.edu.au/view/UQ:87fb7f0>>.

<sup>3</sup> Lorena Allam, 'Pamphlets for voice to parliament could spread misinformation and 'racist messaging', leading yes campaigner says', <https://www.theguardian.com/australia-news/2023/feb/10/pamphlets-for-voice-to-parliament-could-spread-misinformation-and-racist-messaging-leading-yes-campaigner-says>

<sup>4</sup> Axel Bruns, Edward Hurcombe and Stephen Harrington, 'Covering Conspiracy: Approaches to Reporting the COVID/5G Conspiracy Theory' (2022) 10(6) *Digital Journalism* 930; Yochai Benkler, Robert Faris, and Hal Roberts, *Network Propaganda Manipulation, Disinformation, and Radicalization in American Politics* (Oxford University Press, 2018) <<https://academic.oup.com/book/26406>>.

<sup>5</sup> Graham et al (n 1).

<sup>6</sup> Ariadna Matamoros-Fernández, Louisa Bartolo and Luke Troynar, 'Humour as an Online Safety Issue: Exploring Solutions to Help Platforms Better Address This Form of Expression' (2023) 12(1) *Internet Policy Review* <<https://policyreview.info/articles/analysis/humour-as-online-safety-issue-exploring-solutions-social-media-platforms>> ('Humour as an Online Safety Issue').

<sup>7</sup> Timothy Graham and Tobias Keller, 'Bushfires, Bots and Arson Claims: Australia Flung in the Global Disinformation Spotlight' (online, 10 January 2020) <<https://theconversation.com/bushfires-bots-and-arson-claims-australia-flung-in-the-global-disinformation-spotlight-129556>>.

**Recommendation 2:** Alternatively, if the aim is not to broadly ‘combat misinformation and disinformation’, but to instead to promote process and systems change within digital platforms, alongside greater transparency and accountability around their efforts to address content that causes or contributes to ‘serious harm’ (what the substance of the Bill proposes), we recommend refining the scope of the Bill to make this purpose clear from the outset. This could include re-framing the bill as a ‘Digital Platforms Transparency and Accountability Bill’, rather than a Bill that directly ‘combats’ misinformation and disinformation. This provides scope for continued policy reform and activity to address the important contributing factors highlighted above.

### **Removal of the Distinction Between Misinformation and Disinformation**

We recommend removing the distinction between ‘misinformation’ and ‘disinformation’ provided for in clause 7 of the Bill and instead make the central criterion whether the continued circulation of that content could cause or contribute to serious harm, as defined in clause 2 of the Bill. There are several reasons for this recommendation. First, we believe that the distinction is not necessary to or helpful for the purposes of carrying out the objectives of the Bill. Whether the content was shared in good faith or with the intent of deceiving another person is largely irrelevant for addressing the circulation of content that could cause or contribute to serious harm.<sup>8</sup> For example, false information about the effectiveness of COVID-19 vaccines or the integrity of Australia’s electoral system is equally as likely to cause harm to Australian society as false content that is circulated deliberately and in full knowledge of its inaccuracy. Intentionality is important from a national security perspective, as is the case with understanding and mapping foreign adversaries in contested digital spaces (e.g., state-backed election interference), however within the remit of the Bill it is inessential.

Second, there are practical problems with basing a distinction between misinformation and disinformation on whether there is intent to deceive. Notably, it requires an assessment of individual users’ intrinsic motivations for sharing content. This can be time-consuming and exceptionally difficult to do in practice, and at scale, given intrinsic motivations of the actor cannot be done by observing the nature of the content alone. Moreover, the distinction between misinformation and disinformation becomes difficult to determine as content gets reshared on digital platforms. The same piece of content, distributed at first as deliberate *disinformation* by domestic or international influence operations, is likely to be subsequently picked up and shared *misinformation* by ordinary

---

<sup>8</sup> This appears to be based largely on the influential work of Claire Wardle and Hossein Derajkshan, *Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making* (2007) Council of Europe.

social media users deceived by the disinformation campaign. For instance, we note that on the final page of the Bill's guidance "unauthorised electoral or referendum content that is disinformation" is treated differently to "unauthorised electoral or referendum content that is misinformation". Considered in the context of the events around the recent NSW election, where false claims of ballot tampering on social media led to the harassment of poll workers,<sup>9</sup> it is effectively impossible to know whether the people sharing that information are doing so because they want to undermine faith in the election (and know that the electoral commission workers are doing their jobs as normal), or whether they're just innocently misinterpreting what they're seeing (perhaps based off some deliberate disinformation they have encountered previously). This places an unreasonable burden on digital platforms to distinguish between the two types of content.

Finally, we note the use of the use of the two terms in the bill leads to some confusion. While the draft bill introduces the distinction between misinformation and disinformation in clause 7, the two terms consistently appear together throughout the Bill. Indeed, the guidance note admits to using the two terms interchangeably. We also note the exception of the lone mention of disinformation as distinct from misinformation in subclause 35(1)(a) regarding political speech. We believe that this case would be better dealt with explicitly rather than creating conflicting terminology that will be difficult, or impossible to prove.

We recognise that the distinction between misinformation and disinformation will remain relevant with respect to measures against the *actors* who are engaged in spreading such problematic content. In such circumstances, a differentiated response that treats deliberate information actors differently to those social media users who unwittingly fell for such attempts to manipulate public debate and opinion in Australia is appropriate and sensible. However, we believe that action against harmful content (which can be undertaken more quickly) is of higher priority than action against the spreaders of such harmful content. For this reason, we believe that the central criterion must simply be whether the continued circulation of that content could cause or contribute to serious harm, as defined in clause 2 of the Bill.

Similarly, the myopic focus on *content* in the Bill neglects the findings of, and toolkits developed by, nationally significant ARC-funded projects that detect mis- and disinformation through analysis of the

---

<sup>9</sup> Ariel Bogle, 'False Vote-Rigging Claims Force AEC to Ask TikTok to Remove Election Footage', *ABC News* (online, 30 August 2022) <<https://www.abc.net.au/news/science/2022-08-31/aec-staff-tiktok-federal-election-disinformation/101382796>>.

coordinated social context in which it occurs.<sup>10</sup> To infer the intentionality of a borderline piece of content, disinformation researchers increasingly focus on its sharing dynamics in online communities – that is, the social and participatory context in which it circulates. A piece of content may not need to be explicitly false to cause harm, if it is repeatedly shared by hundreds of accounts in a short time window. For example, during the 2019/20 bushfires, QUT DMRC researchers found that a network of activists and fake accounts managed to mislead the Australian public by getting the hashtag #ArsonEmergency to trend on Twitter and attract attention of mass media.<sup>11</sup> It is also evident that coordinated activity by foreign adversaries in contested digital spaces does not require the sharing of explicitly false content - it is the volume and reach of decontextualised content, shared at scale with the intent to co-opt organic online communities – that matters, as is the case with Russian government war propaganda about the Ukraine war.<sup>12</sup> The bill, as it stands, neglects *coordinated behaviour*, and this needs to be addressed by evaluating the context of content rather than just the content itself.

**Recommendation 3:** To the extent that the aim of the bill is to encourage the sharpening of measures against the circulation of *content* that is likely to cause harm, we recommend that the distinction between the two terms be removed altogether. Instead, we suggest it be replaced with a clarification that for the purposes of the bill the intent behind the potentially harmful content is irrelevant. A local exception can be made in clause 35(1)(a) for political speech if the focus on disinformation alone in that clause was indeed intended. We note that the removal of a distinction between misinformation and disinformation in the Bill does not preclude the voluntary or mandatory industry codes and standards required by the bill from considering intent in the definition of measures to be taken against harmful content and, especially, the actors who disseminate such content. The removal of any test of intent also has the added benefit of potentially increasing the speed of responses to harmful content, which will be important for problematic information related to acute crises and current events.

---

<sup>10</sup> Graham et al (n 1).

<sup>11</sup> Tobias Keller et al, 'Coordinated Inauthentic Behaviour and Other Influence Operations in Social Media Spaces' [2020] *AoIR Selected Papers of Internet Research* <<https://spir.aoir.org/ojs/index.php/spir/article/view/11132>>.

<sup>12</sup> Timothy Graham and Jay Daniel Thompson, 'Russian Government Accounts Are Using a Twitter Loophole to Spread Disinformation', *The Conversation* (online, 15 March 2022) <<http://theconversation.com/russian-government-accounts-are-using-a-twitter-loophole-to-spread-disinformation-178001>>.

## Tightening the Definition of ‘good faith’ to Address the Influence of Satire in Spreading Misinformation and Disinformation

The Bill, as it is currently drafted, excludes all entertainment, parody and satire as content for disinformation purposes if it is said or done in “good faith”. However, we argue that the definition of “good faith” content requires tightening because humour can be used as a catch-all to legitimise highly problematic and harmful content online. While satire and parody content are widely accepted in Australia, particularly when said or done in “good faith”, it is still necessary to exercise caution and diligence to avoid or at least minimise harmful consequences.<sup>13</sup> For this reason, we suggest including satire and parody as content that has the potential to harm, especially when this satire and parody promotes racist and other discriminatory narratives.

Humour and its diverse rhetorical devices, including satire and parody, can be harmful,<sup>14</sup> especially online where context is easily lost and intent is difficult to assess.<sup>15</sup> On digital platforms, not only can ‘bad actors’ weaponise satire and parody to spread racial hate and dangerous narratives,<sup>16</sup> ordinary people can also advertently and inadvertently harm through their everyday engagements with humorous expression.<sup>17</sup> Although some digital platforms are trialling content moderation tools that can add context to misleading satire, such as the Twitter Community Notes and Facebook adding labels to “satire pages”, the effectiveness of these type of tools without proper oversight is dubious.<sup>18</sup>

Within the context of disinformation online, fact-checking organisation and scholars have raised the alarm about how satire is commonly used to excuse disinformation,<sup>19</sup> noting that this practice is

---

<sup>13</sup> Judith Bannister, ‘It’s Not What You Say But the Way That You Say It: Australian Hate Speech Laws and the Exemption of “Reasonable” Expression’ (2008) 36(1) *Florida State University Law Review* 23.

<sup>14</sup> Katharine Gelber, ‘Differentiating Hate Speech: A Systemic Discrimination Approach’ (2021) 24(4) *Critical Review of International Social and Political Philosophy* 393; Sharon Lockyer and Michael Pickering, ‘Introduction: The Ethics and Aesthetics of Humour and Comedy’ in Sharon Lockyer and Michael Pickering (eds), *Beyond a Joke: The Limits of Humour* (Palgrave Macmillan UK, 2005) 1 <[https://doi.org/10.1057/9780230236776\\_1](https://doi.org/10.1057/9780230236776_1)>; Simon Weaver, *The Rhetoric of Racist Humour: US, UK and Global Race Joking* (Routledge, 2016).

<sup>15</sup> Matamoros-Fernández, Bartolo and Troynar (n 5).

<sup>16</sup> Viveca S Greene, “‘Deplorable’ Satire: Alt-Right Memes, White Genocide Tweets, and Redpilling Normies’ (2019) 5(1) *Studies in American Humor* 31.

<sup>17</sup> Ariadna Matamoros-Fernández, Aleesha Rodriguez and Patrik Wikström, ‘Humor That Harms? Examining Racist Audio-Visual Memetic Media on TikTok During Covid-19’ (2022) 10(2) *Media and Communication* 180.

<sup>18</sup> Sameer Borwankar, Jinyang Zheng and Karthik Natarajan Kannan, ‘Democratization of Misinformation Monitoring: The Impact of Twitter’s Birdwatch Program’ (SSRN Scholarly Paper No 4236756, 3 October 2022) <<https://papers.ssrn.com/abstract=4236756>>.

<sup>19</sup> AFP, ‘It’s No Joke: Across Globe, Satire Morphs into Misinformation’, *France 24* (online, 15 December 2022) <<https://www.france24.com/en/live-news/20221215-it-s-no-joke-across-globe-satire-morphs-into-misinformation>>; First Draft, “‘It Was Just a Joke’: How Satire Is Used to Excuse Disinformation in Elections”, *First Draft* (online, 12 November 2019) <<https://firstdraftnews.org/articles/it-was-just-a-joke-how-satire-is-used-to-excuse-disinformation-in-elections/>>; Claire Wardle, ‘The Need for Smarter Definitions and Practical, Timely Empirical Research on Information Disorder’ (2018) 6(8) *Digital Journalism* 951.

especially dangerous when satirical disinformation is targeted at historically marginalised groups and has racist and other discriminatory overtones.<sup>20</sup> Additionally, political satire and humour can also be employed to circumvent conventional communication regulations under the guise of ‘just a joke’, ‘satirical’ or ‘lighthearted’ content.<sup>21</sup> This type of content can be deliberately weaponised to promote harmful narratives without explicitly stating them. A key example of this is the animated political satire web series “Pauline Hanson’s Please Explain” published on Senator Pauline Hanson’s official YouTube channel which, at time of writing, has 69,000 subscribers.<sup>22</sup> It frequently presents exaggerated and unfounded narratives about senior politicians to bolster Senator Hanson’s political stance. In one of the animated cartoons – now deleted after it received public backlash – Senator Penny Wong, an Asian-Australian politician, is portrayed as malevolently giving Prime Minister Anthony Albanese soup containing a bat, alluding to the debunked theory linking COVID-19’s origins to bat consumption. This connection perpetuates racial biases, especially given that such theories, when initially circulated during the pandemic, were associated with increased racially motivated attacks on Asian-Australians.<sup>23</sup>

**Recommendation 4:** We recommend that satire and parody be treated with care in any regulation targeting online disinformation and its potential to harm. The Bill, as it is currently drafted, excludes all entertainment, parody and satire as content for disinformation purposes if done in “good faith”. Regarding the “good faith” provision, legal scholars in Australia have noted that exemptions of content performed in “good faith” in laws targeted at protecting vulnerable communities can in fact further harm those that these laws try to protect in the first place.<sup>24</sup> We suggest that the definition of “good faith” does not apply to baseless and far-reaching claims that, for example, perpetuate racist tropes or damage the credibility of Australia’s electoral process. Additionally, as previously mentioned, such “good faith” intent is almost impossible to assess in the online context.<sup>25</sup> Second, we note that the current wording of this exception in the Bill assumes that satire and parody do not pose a risk within disinformation disorders, but there is ample evidence about how humour can in

---

<sup>20</sup> Elise Worthington, Ariel Bogle and Michael Workman, ‘Fringe Parties Spreading Trump-Style “big Lie” Conspiracies Are Challenging the Integrity of Australia’s Electoral System’, *ABC News* (online, 16 May 2022) <<https://www.abc.net.au/news/2022-05-17/australians-spreading-trump-election-conspiracies/101057226>>.

<sup>21</sup> First Draft (n 18).

<sup>22</sup> *A New Term | Pauline Hanson’s Please Explain* (Directed by Pauline Hanson’s Please Explain, 3 February 2023) <<https://www.youtube.com/watch?v=-YOeXKtvshk>>.

<sup>23</sup> Naaman Zhou, ‘Asian Australians Threatened and Spat on in Racist Incidents amid Coronavirus | Australia News | The Guardian’, *The Guardian* (online, 24 July 2020) <<https://www.theguardian.com/australia-news/2020/jul/24/asian-australians-threatened-and-spat-on-in-racist-incidents-amid-coronavirus>>.

<sup>24</sup> Judith Bannister (n 12).

<sup>25</sup> Whitney Phillips and Ryan Milner, *The Ambivalent Internet: Mischief, Oddity, and Antagonism Online* (Polity, 2017).



fact harm in information crises, such as COVID-19 and Australian elections.<sup>26</sup> Therefore, we recommend a reconsideration of the potentially harmful effects of humour in this Bill and suggest that satire and parody is considered content that has potential to harm.

### **News Media Exclusions**

With respect to the various carve-outs for ‘professional news content’, we note that this bill potentially excludes accounts that produce news-like content but consistently fail to adhere to or enforce robust editorial standards and ethics in their reporting. This includes those that delegitimise the science of anthropogenic climate change, or question the legitimacy of electoral processes without evidence, or promote anti-vaccination sentiment.

Though the carve-out for ‘professional news content’ likely exists because the activity of such outlets will presumably be captured through existing mechanisms (i.e. existing ACMA powers), the risk is that a host of news producers – ones which perhaps style themselves as ‘legitimate’, ‘truthful’ news outlets, but which ultimately don’t follow robust editorial standards may slip through the cracks between the two enforcement domains. Examples here include hyper-partisan news programs and podcasts on social media platforms such as *YouTube*, many of which were responsible for some of the most egregious examples of mis- and disinformation during the height of the COVID-19 pandemic. Such outlets would not be captured by this bill, but also cannot be trusted to act professionally and responsibly of their own accord.

Furthermore, that the bill is focussed entirely on digital platforms, and excludes these ‘professional’ news outlets, reveals a mis-match between the overall aims of the legislation and the reality of how mis- and disinformation moves through our culture, and thus does not properly address the likely cases of greatest harm. As numerous studies have shown (a good example being Tsifati, et al., 2020)<sup>27</sup>, most disinformation remains relatively obscure to – and goes unnoticed by – the public if it is shared solely on digital platforms (often being largely limited to, for instance, existing conspiracy-minded or hyper-partisan groups). It often ‘breaks out’ to a wider audience, however, only once that disinformation is reported on by well-meaning journalists working for legitimate (and often highly trusted) news organisations, or less-scrupulous news workers who may promote this work to support

---

<sup>26</sup> Matamoros-Fernández, Rodríguez and Wikström (n 16); Elise Worthington, Ariel Bogle and Michael Workman (n 19).

<sup>27</sup> Tsifati, Y., Boomgaarden, H. G., Strömbäck, J., et al. (2020). Causes and consequences of mainstream media dissemination of fake news: literature review and synthesis. *Annals of the International Communication Association*, 44(2), pp. 157-173.

the partisan goals of their masthead. Work undertaken previously by members of the DMRC,<sup>28</sup> for instance, showed that the entirely erroneous link between the onset and impact of COVID-19 and the rollout of 5G mobile technology, only really entered the broader public consciousness once it had been discussed by celebrity figures, and then reported on by journalists as part of their celebrity news coverage. Prior to that point, it had had little impact.

Though we acknowledge that it is beyond the current scope of this bill, and is an aspirational comment only, a far more powerful approach would include both digital platforms *and* professional news outlets, ensuring that potentially harmful content has limited impact with the wider public, and isn't unwittingly amplified to a dramatically larger audience than it would have otherwise enjoyed.

The debates related to the EU Digital Services Act (EU DSA) are directly relevant to this issue because amongst its various provisions, the EU DSA imposes obligations on the largest platforms (>=45million users in the EU) to tackle 'systemic risks' - which include disinformation. Platforms are expected to respect "fundamental rights" in the process – and included within the list of fundamental rights is "media freedom and pluralism".<sup>29</sup> In other words, the EU DSA very much captures this tension between regulating to tackle disinformation and upholding protections for media freedom. Related to this discussion we also note that the Bill has no counterweight measures or clauses to prevent over-restriction of content by platforms. While the EU DSA required respecting 'fundamental rights', this Bill has no such expectation or requirement that platforms don't go too far in the restriction of content.

In the initial draft of the EU DSA, news media/journalists were not mentioned at all.<sup>30</sup> Post-draft consultation saw various actors, including news media organisations, lobby for a "non-interference principle" to be included in the EU DSA. As Papaevangelou notes,<sup>31</sup> the proposed "non-interference principle" was meant to include 3 things:

1. "A restriction of platforms to meddle with editorial content"

---

<sup>28</sup> Bruns, Hurcombe and Harrington (n 3); Axel Bruns, Stephen Harrington and Edward Hurcombe, 'Coronavirus Conspiracy Theories: Tracing Misinformation Trajectories from the Fringes to the Mainstream' in Monique Lewis, Eliza Govender and Kate Holland (eds), *Communicating COVID-19: Interdisciplinary Perspectives* (Springer International Publishing, 2021) 229 <[https://doi.org/10.1007/978-3-030-79735-5\\_12](https://doi.org/10.1007/978-3-030-79735-5_12)>.

<sup>29</sup> Charis Papaevangelou, "'The Non-Interference Principle': Debating Online Platforms' Treatment of Editorial Content in the European Union's Digital Services Act' [2023] *European Journal of Communication* 02673231231189036.

<sup>30</sup> Doris Buijs, 'The Digital Services Act & the Implications for the Safety of Journalists - DSA Observatory', *DSA Observatory* (27 October 2022) <<https://dsa-observatory.eu/2022/10/27/the-digital-services-act-the-implications-for-the-safety-of-journalists-part-2/>>.

<sup>31</sup> Papaevangelou (n 27).

2. “Establishment of dedicated communication channels between platforms and news organisations to rectify wrongful content restrictions and be informed a priori of any changes in their algorithm”
3. “Integration of the respect for media freedom and pluralism in platforms’ terms and conditions”

However, this was ultimately unsuccessful. Instead, as mentioned above, what was introduced was a provision which requires platforms to uphold media pluralism and media freedom as part of a commitment to fundamental rights. A key concern throughout was how to define “media” - and the concern that a broad definition would give cover to all sorts of bad actors operating as “journalists.”

Much remains to be seen as the EU DSA is implemented, but the relevant point perhaps is that some argue that as part of a “positive obligation” to address disinformation, platforms may focus on “promoting diverse content and authoritative journalism on matters of public interest” - which is an approach also favoured in the EU’s Code of Practice on Disinformation.<sup>32</sup>

Importantly, the EU is also discussing a proposed European Media Freedom Act (EMFA) which will have major implications for how digital platforms treat media providers more generally, with various privileges being considered (Article 17) (in practice, these privileges would mean commercial digital platforms would face greater barriers for removing content by media as opposed to UGC). Here again the question of how to define “media” has been especially sticky – with some scholars urging the EU to look to international and European human rights law to avoid overly-broad or incoherent definitions.<sup>33</sup> Civil Society groups like the Electronic Frontiers Foundation argue that it would make disinformation issues worse, because platforms will be prevented from speedily dealing with disinformation stemming from self-declared “media” entities.<sup>34</sup>

**Recommendation 5:** We recommend a careful examination of the definition of a “news organisation” and consider how exclusions for “professional news content” could be exploited by misinformation agents who self-identify as “journalists”.

---

<sup>32</sup> Doris Buijs (n 28).

<sup>33</sup> Joan Barata, ‘Protecting Media Content on Social Media Platforms: The European Media Freedom Act’s Biased Approach’ [2022] *Verfassungsblog* <<https://verfassungsblog.de/emfa-dsa/>>.

<sup>34</sup> David Greene, Paige Collings, and Christoph Schmon, ‘Online Platforms Should Stop Partnering with Government Agencies to Remove Content’, *Electronic Frontier Foundation* (online, 12 August 2022) <<https://www.eff.org/deeplinks/2022/08/online-platforms-should-stop-partnering-government-agencies-remove-content>>.

## Information-gathering powers

We support the intent regarding ACMA being granted powers to obtain information required to make assessments regarding the efficacy of platform-based solutions to the problems outlined in the bill. We recommend an expansion of these powers to also include requirements on Digital Platforms to offer more equitable and wholistic data-sharing for external analysis by academics, civil society and/or governments. While many platforms currently provide limited data access arrangements, these are often piecemeal, voluntary, and prone to significant change without notice.<sup>35</sup> The example of X (formerly Twitter) recently discontinuing its Academic Application Programming Interface (API) and replacing its remaining APIs with prohibitively expensive subscription-based data-access models has now removed one of the most historically important mechanisms available to academics and civil society actors to track misinformation on the platform. The DMRC has been at the forefront of research to map and expose significant misinformation and coordinated inauthentic activity on Twitter,<sup>36</sup> and the removal of this data access has now significantly hampered our ability to track this and other significantly problematic activity on this platform.

Access to data is doubly important as it provides not only a means through which researchers can continue to monitor and track misinformation activity, but a means through which researchers can refine or develop new methodologies to keep pace with, and counter, emerging tactics and technologies used by misinformation agents.

Digital platforms may refer to specific institutional arrangements such as Social Science One, or 'transparency dashboards' like Meta and Alphabet's advertising transparency initiatives, as mechanisms through which they already provide such data access. These data access programs and dashboards are however severely limited in terms of providing meaningful data to allow the review of measures implemented by digital platform services to prevent or respond to misinformation or disinformation on their services. Many key measurements and data provided through these services lack provenance, and there have been significant cases where datasets provided have contained egregious errors or omissions.<sup>37</sup> Data quality aside, these services often come with restrictions on use of data by approved researchers, or are significantly limited in terms of what

---

<sup>35</sup> Axel Bruns, 'After the "APIcalypse": Social Media Platforms and Their Fight against Critical Scholarly Research' (2019) 22(11) *Information, Communication & Society* 1544 ('After the "APIcalypse"').

<sup>36</sup> Frances Shaw et al, 'Sharing News, Making Sense, Saying Thanks' (2013) 40(1) *Australian Journal of Communication* 23; Timothy Graham et al, *Like a Virus: The Coordinated Spread of Coronavirus Disinformation* (Report, Centre for Responsible Technology, 1 June 2020) <<https://apo.org.au/node/305864>>.

<sup>37</sup> Craig Timberg, 'Facebook Made Big Mistake in Data It Provided to Researchers, Undermining Academic Work', *Washington Post* (online, 11 September 2021) <<https://www.washingtonpost.com/technology/2021/09/10/facebook-error-data-social-scientists/>>.

content they index and include.

While we argue for widened scope regarding data access provision of public data, we support provisions limiting accessibility of private content such as private messages. We would also support any measures to ensure that such data access falls under the scope of existing national ethics guidelines.

**Recommendation 6:** We recommend strengthening of the information gathering powers to include requirements on Digital Platforms to offer more equitable and wholistic data-sharing for external analysis by academics, civil society and/or governments. This would include safeguards against the sharing of private data, as already outlined in the bill.

## References

*A New Term* | *Pauline Hanson's Please Explain* (Directed by Pauline Hanson's Please Explain, 3 February 2023) <<https://www.youtube.com/watch?v=-YOeXKtvshk>>

AFP, 'It's No Joke: Across Globe, Satire Morphs into Misinformation', *France 24* (online, 15 December 2022) <<https://www.france24.com/en/live-news/20221215-it-s-no-joke-across-globe-satire-morphs-into-misinformation>>

Ariel Bogle, 'False Vote-Rigging Claims Force AEC to Ask TikTok to Remove Election Footage', *ABC News* (online, 30 August 2022) <<https://www.abc.net.au/news/science/2022-08-31/aec-staff-tiktok-federal-election-disinformation/101382796>>

Barata, Joan, 'Protecting Media Content on Social Media Platforms: The European Media Freedom Act's Biased Approach' [2022] *Verfassungsblog* <<https://verfassungsblog.de/emfa-dsa/>>

Borwankar, Sameer, Jinyang Zheng and Karthik Natarajan Kannan, 'Democratization of Misinformation Monitoring: The Impact of Twitter's Birdwatch Program' (SSRN Scholarly Paper No 4236756, 3 October 2022) <<https://papers.ssrn.com/abstract=4236756>>

Bruns, Axel, 'After the "APIcalypse": Social Media Platforms and Their Fight against Critical Scholarly Research' (2019) 22(11) *Information, Communication & Society* 1544

Bruns, Axel, Stephen Harrington and Edward Hurcombe, 'Coronavirus Conspiracy Theories:

Tracing Misinformation Trajectories from the Fringes to the Mainstream' in Monique Lewis, Eliza Govender and Kate Holland (eds), *Communicating COVID-19: Interdisciplinary Perspectives* (Springer International Publishing, 2021) 229 <[https://doi.org/10.1007/978-3-030-79735-5\\_12](https://doi.org/10.1007/978-3-030-79735-5_12)>

Bruns, Axel, Edward Hurcombe and Stephen Harrington, 'Covering Conspiracy: Approaches to Reporting the COVID/5G Conspiracy Theory' (2022) 10(6) *Digital Journalism* 930

Doris Buijs, 'The Digital Services Act & the Implications for the Safety of Journalists (Part 1)', *DSA Observatory* (27 October 2022) <<https://dsa-observatory.eu/2022/09/29/digital-services-act-implications-for-news-media-journalistic-content-part-1/>>

David Greene, Paige Collings, and Christoph Schmon, 'Online Platforms Should Stop Partnering with Government Agencies to Remove Content', *Electronic Frontier Foundation* (online, 12 August 2022) <<https://www.eff.org/deeplinks/2022/08/online-platforms-should-stop-partnering-government-agencies-remove-content>>

Elise Worthington, Ariel Bogle and Michael Workman, 'Fringe Parties Spreading Trump-Style "big Lie" Conspiracies Are Challenging the Integrity of Australia's Electoral System', *ABC News* (online, 16 May 2022) <<https://www.abc.net.au/news/2022-05-17/australians-spreading-trump-election-conspiracies/101057226>>

First Draft, "It Was Just a Joke": How Satire Is Used to Excuse Disinformation in Elections', *First Draft* (online, 12 November 2019) <<https://firstdraftnews.org/articles/it-was-just-a-joke-how-satire-is-used-to-excuse-disinformation-in-elections/>>

Fredericks, Bronwyn et al, 'Burden of the Beast : Countering Conspiracies and Misinformation within Indigenous Communities in Australia' (2022) 25(1) *M/C Journal* <<https://espace.library.uq.edu.au/view/UQ:87fb7f0>>

Gelber, Katharine, 'Differentiating Hate Speech: A Systemic Discrimination Approach' (2021) 24(4) *Critical Review of International Social and Political Philosophy* 393

Graham, Timothy et al, '#IStandWithDan versus #DictatorDan: The Polarised Dynamics of Twitter Discussions about Victoria's COVID-19 Restrictions' (2021) 179(1) *Media International Australia* 127

Graham, Timothy et al, *Like a Virus: The Coordinated Spread of Coronavirus Disinformation* (Report, Centre for Responsible Technology, 1 June 2020) <<https://apo.org.au/node/305864>>

Greene, Viveca S, “‘Deplorable’ Satire: Alt-Right Memes, White Genocide Tweets, and Redpilling Normies’ (2019) 5(1) *Studies in American Humor* 31

Judith Bannister, ‘It’s Not What You Say But the Way That You Say It: Australian Hate Speech Laws and the Exemption of “Reasonable” Expression’ (2008) 36(1) *Florida State University Law Review* 23

Keller, Tobias et al, ‘Coordinated Inauthentic Behaviour and Other Influence Operations in Social Media Spaces’ [2020] *AoIR Selected Papers of Internet Research*  
<<https://spir.aoir.org/ojs/index.php/spir/article/view/11132>>

Lockyer, Sharon and Michael Pickering, ‘Introduction: The Ethics and Aesthetics of Humour and Comedy’ in Sharon Lockyer and Michael Pickering (eds), *Beyond a Joke: The Limits of Humour* (Palgrave Macmillan UK, 2005) 1 <[https://doi.org/10.1057/9780230236776\\_1](https://doi.org/10.1057/9780230236776_1)>

Matamoros-Fernández, Ariadna, Louisa Bartolo and Luke Troynar, ‘Humour as an Online Safety Issue: Exploring Solutions to Help Platforms Better Address This Form of Expression’ (2023) 12(1) *Internet Policy Review* <<https://policyreview.info/articles/analysis/humour-as-online-safety-issue-exploring-solutions-social-media-platforms>>

Matamoros-Fernández, Ariadna, Aleesha Rodriguez and Patrik Wikström, ‘Humor That Harms? Examining Racist Audio-Visual Memetic Media on TikTok During Covid-19’ (2022) 10(2) *Media and Communication* 180

Naaman Zhou, ‘Asian Australians Threatened and Spat on in Racist Incidents amid Coronavirus | Australia News | The Guardian’, *The Guardian* (online, 24 July 2020)  
<<https://www.theguardian.com/australia-news/2020/jul/24/asian-australians-threatened-and-spat-on-in-racist-incident-amid-coronavirus>>

Papaevangelou, Charis, “‘The Non-Interference Principle’: Debating Online Platforms’ Treatment of Editorial Content in the European Union’s Digital Services Act’ [2023] *European Journal of Communication* 02673231231189036

Shaw, Frances et al, ‘Sharing News, Making Sense, Saying Thanks’ (2013) 40(1) *Australian Journal of Communication* 23

Timberg, Craig, ‘Facebook Made Big Mistake in Data It Provided to Researchers, Undermining Academic Work’, *Washington Post* (online, 11 September 2021)

<<https://www.washingtonpost.com/technology/2021/09/10/facebook-error-data-social-scientists/>>

Timothy Graham and Jay Daniel Thompson, 'Russian Government Accounts Are Using a Twitter Loophole to Spread Disinformation', *The Conversation* (online, 15 March 2022)

<<http://theconversation.com/russian-government-accounts-are-using-a-twitter-loophole-to-spread-disinformation-178001>>

Timothy Graham and Tobias Keller, 'Bushfires, Bots and Arson Claims: Australia Flung in the Global Disinformation Spotlight' (online, 10 January 2020) <<https://theconversation.com/bushfires-bots-and-arson-claims-australia-flung-in-the-global-disinformation-spotlight-129556>>

Tsfati, Y., Boomgaarden, H. G., Strömbäck, J., et al. Causes and consequences of mainstream media dissemination of fake news: literature review and synthesis (2020) 44(2) *Annals of the International Communication Association* 157-173

Wardle, Claire, 'The Need for Smarter Definitions and Practical, Timely Empirical Research on Information Disorder' (2018) 6(8) *Digital Journalism* 951

Weaver, Simon, *The Rhetoric of Racist Humour: US, UK and Global Race Joking* (Routledge, 2016)

Whitney Phillips and Ryan Milner, *The Ambivalent Internet: Mischief, Oddity, and Antagonism Online* (Polity, 2017)

Yochai Benkler, Robert Faris, and Hal Roberts, *Network Propaganda Manipulation, Disinformation, and Radicalization in American Politics* (Oxford University Press, 2018)

<<https://academic.oup.com/book/26406>>