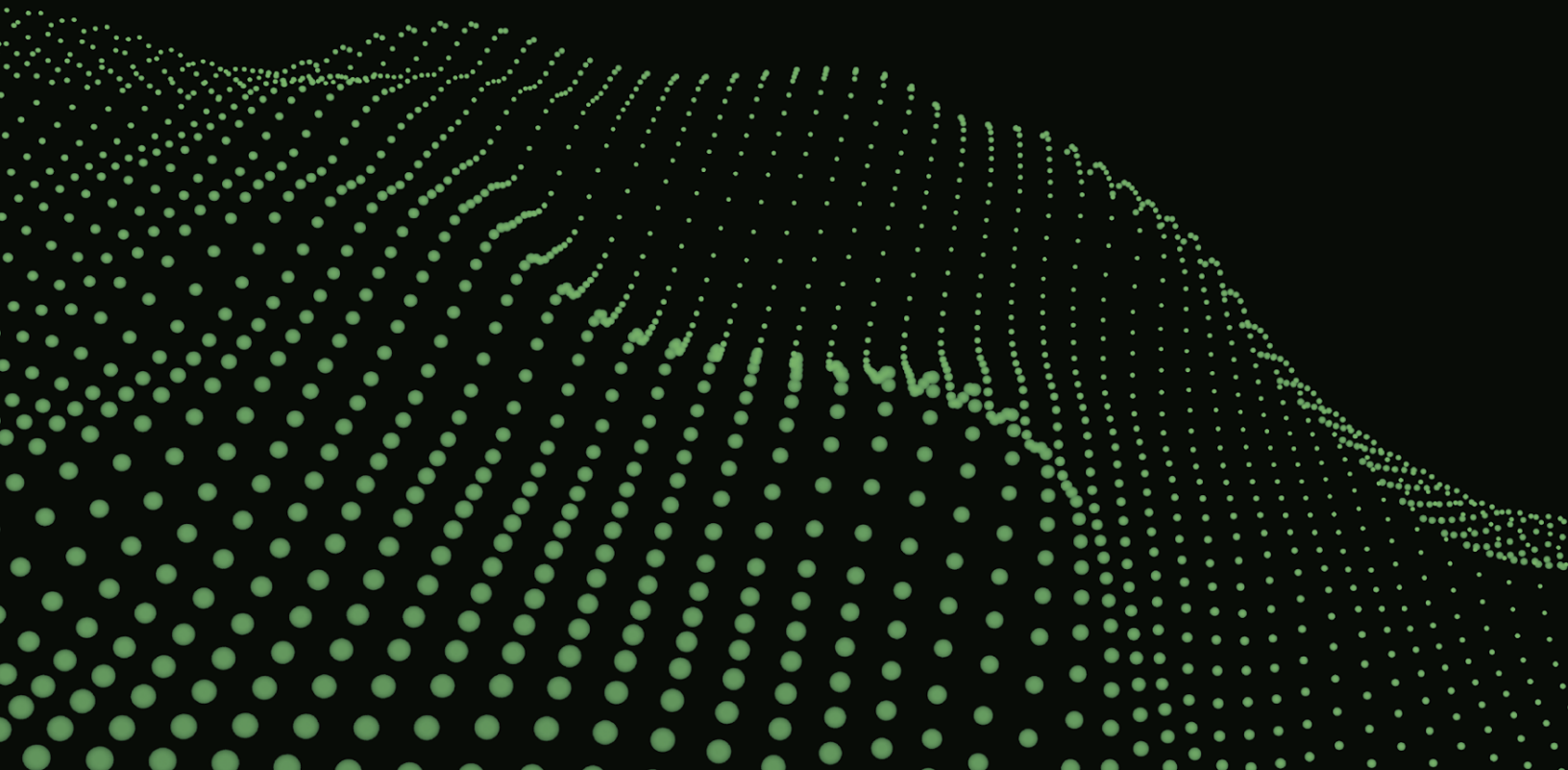# Australian Code of Practice on Disinformation and Misinformation |
Annual Report

Published May 29, 2023

Background

DIGI is a non-profit industry association that advocates for the interests of the digital industry in Australia. DIGI's founding members are Apple, eBay, Google, Linktree, Meta, Spotify, Snap, TikTok, Twitter and Yahoo, and its associate members are Change.org, Gofundme, ProductReview.com.au and Redbubble. DIGI's vision is a thriving Australian digitally-enabled economy that fosters innovation, a growing selection of digital products and services, and where online safety and privacy are protected. DIGI is a key Government partner in efforts to address online harms, data and consumer protection online and to grow the digital economy, through code development, partnerships and advocacy for effective and implementable approaches to technology policy.

The Australian Code of Practice on Disinformation and Misinformation (ACPDM) was developed in response to Australian Government policy announced in December 2019, in response to the ACCC Digital Platforms Inquiry, where the digital industry was asked to develop a voluntary code of practice on disinformation.

DIGI developed the ACPDM with assistance from the University of Technology Sydney's Centre for Media Transition, and First Draft, a global organisation that specialises in helping societies overcome false and misleading information.

The ACPDM was launched in February 2021 and its signatories are Apple, Adobe, Google, Meta, Microsoft, Redbubble, TikTok and Twitter. The ACDM was further strengthened in December 2022 in response to stakeholder feedback received through a planned review of the code that included a six week public consultation.These updates are the latest set in a series of improvements driven by DIGI and code signatories since the code was introduced in February 2021.

Mandatory code commitments include publishing & implementing policies on misinformation and disinformation, providing users with a way to report content against those policies and implementing a range of scalable measures that reduce its spread & visibility (Mandatory commitment #1). Every signatory has agreed to annual transparency reports about those efforts to improve understanding of both the management and scale of mis- and disinformation in Australia (Mandatory commitment #7).

Additionally, there are a series of opt-in commitments that platforms adopt if relevant to their business model: (Commitment #2) Addressing disinformation in paid content; (#3) addressing fake bots and accounts; (#4) transparency about source of content in news and factual information (e.g. promotion of media literacy, partnerships with fact-checkers) and (#5) political advertising; and (#6) partnering with universities/researchers to improve understanding of mis and disinformation.

DIGI has produced this annual report as part of its governance of the ACPDM.

Australian Code of Practice on Disinformation and Misinformation | Annual Report

Published May 29, 2023

Managing Director's statement

By: Sunita Bose
Managing Director, DIGI

The release of our second annual report, and the third set of signatories' transparency reports, under *The Australian Code of Practice on Disinformation and Misinformation* (ACPDM) provides a moment to reflect on the code's evolution since it was launched in February 2021. The code is now no longer new, but it is *novel*, as one of only two such codes in the world. Signatory technology companies have committed to, and continue to evolve, safeguards that protect against online disinformation and misinformation.

A novel code and a highly complex issue means the policy pathway forward is not always clearly laid out. DIGI has had to develop bespoke structures to deepen accountability for continued industry action in efforts to address mis-and disinformation. In addition to the independent members on our code committees, we have continued to engage Hal Crawford, our independent reviewer, who develops and updates best practice reporting guidelines to drive improvements in signatories' transparency reports. He assesses each transparency report against those guidelines, asks signatories for improvements, and attests claims prior to publication. The latest version of these guidelines as well the collective independent assessment of the reports, is included in the appendix of this report. We thank all of our independent members for their continued support.

While the first set of transparency reports, released in May 2021, contained mostly qualitative descriptions of signatories' efforts, this independent process has seen subsequent reports contain more quantitative insights, laying the foundations for comparisons. TikTok has seen improvements in its 'proactive removal' rates of misinformation content – its enforcement of content prior to users even *viewing* it increased from 37.6% in Q1 to 69.8% in Q4 of 2022. In Q4, 89.1% of harmful misinformation was removed before a user *reporting* it. Meta has reported on its work to display warnings on content that is

found to be false by independent third-party fact checkers. Over 2022, they displayed warnings on over 9 million distinct pieces of content in Australia.

Yet in their efforts toward transparency, signatories have to be extremely careful that the information they release does not inadvertently release the playbook to amplify harmful misinformation, as bad actors are relentless and resourceful. Google's report details a Chinese influence campaign known as 'dragonbridge' that perpetuates disinformation on news topics such as China's COVID-19 response to the war in Ukraine. Despite this content receiving relatively low engagement from users, in 2022 Google disrupted over 50,000 instances of 'dragonbridge' activity across its services, and terminated over 100 000 accounts.

These case studies show the shifts in issues prone to disinformation. While efforts continue in relation to COVID-19, 2022 presented new challenges that signatories have reported upon, such as the war in Ukraine. Microsoft reported that Russian influence operations leveraged energy and food supply concerns across Europe to weaken the alliance against the Russian invasion.

Importantly, this was the first Australian federal election since the code was introduced, and we believe the code positively contributed to the Australian Electoral Commission's determination that this election saw much lower levels of electoral mis and disinformation this election than in other like minded democratic elections across the globe.[1] The extensive electoral integrity work that signatories undertook to achieve that outcome is detailed in the report, with multiple signatories directing Australians to the AEC to elevate official electoral information, in addition to content removal. We know that signatories are now preparing similar work to prevent mis- and disinformation in relation to the referendum into an Aboriginal and Torres Strait Islander Voice to Parliament, which we expect will be reported upon in their 2023 calendar year reports.

Content removal is often the focus when it comes to conversation about how we tackle misinformation, but it is only one lever in how we effectively respond to the challenge holistically. This is partly due to the diversity of business models and services among signatories, and a recognition that research and education measures are essential to a whole of society response'. The reports show how signatories are supporting important media and digital literacy initiatives both 'on platform', such as through providing fact checking training to influencers and helping people identify trusted information, and 'off platform' to sponsor important expert-led programs.

DIGI is committed to ensuring that the code helps shed light on the scale and management of mis- and disinformation on digital products and services. Since our last annual report, DIGI strengthened *The Australian Code of Practice on Misinformation and Disinformation*, making changes in response to stakeholder feedback received through a planned review of the code, that are detailed in this report.

We welcome the announcement by the Government about the ACMA receiving greater oversight powers of the code, which will reinforce DIGI's efforts and formalise our long-term working relationship with the regulator in relation to combatting misinformation online. Sustained shifts in the fight against mis- and disinformation rely on concerted efforts from industry, Governments, civil society and the community, and we're committed to continuing to play our part.

Sunita Bose
Managing Director, DIGI

---

[1] Australian Electoral Commission, Press Release [03/08/2022], *AEC celebrates successful disinformation partnerships*,

Part 1 | **Outcome of 2022 Review of the ACPDM**

On June 6, 2022, the signatories to the ACPDM commenced a review of the code, as required by section 7.8, accompanied by a DIGI Discussion Paper to assist public consultation that provided background and eight specific questions and proposals for consideration, on which stakeholder views were encouraged:[2]

1) *Should the ACPDM cover a broader scope of signatories? If so should:*
   a) *the ACMA have a continued role in identifying those services that are within the scope of the Code; and*
   b) *what should the criteria be for a company's eligibility to participate in the Code?*
2) *Should the ACPDM take an opt-out rather than an opt-in approach to the optional commitments under the Code?[3]*
3) *Should the definition of harm be amended to deal with concerns about the narrowness of 'serious and imminent threat' language?[4]*
4) *Should the exemption for professional news content be revised so that it is clearer? Should the Code be extended to cover news aggregation services?[5]*
5) *Should the Code be extended to include private messaging services?[6]*
6) *Should the approach of the Code to issues-based advertising be clarified?[7]*
7) *Should the Code better define paid and sponsored content?[8]*
8) *Is the Code meeting the needs of industry and the community to balance concerns about misinformation and disinformation with the need to protect freedom of expression online?*

The ACPDM was updated on 22 December 2022 to reflect the outcome of the review. In addition to the 13 submissions received, the review process took into account the ACMA's June 2021 *Report to government on the adequacy of digital platforms disinformation and news quality measures* (Report to Government).[9] It should be noted that the governance model for the ACPDM was introduced in October 2021, after the ACMA conducted its formal review of the ACPDM, and its *Report to Government* was published in March 2022. Therefore, the 2022 Code review did not take did not take into account the Code's governance arrangements. The launch of the 2022 consultation on the ACPDM also pre-dated the EU Strengthened Code of Practice on Disinformation (EU Strengthened Code), released on June 16, 2022[10]. However, some submissions contained feedback that referenced developments concerning the EU Strengthened Code, and therefore the relevant changes to the EU Strengthened Codes were also considered.

A detailed response to submissions, and the specific questions in the DIGI Discussion Paper[11] was published on the DIGI website on 22 December 2022, together with the updated ACPDM. Some of the key changes made as part of the review include:

---

[2] DIGI, *ACPDM, 2022 Review of Australian Code of Practice on Disinformation and Misinformation, Discussion Paper*, June 6 2022
[3] *ACMA Report to Government*, findings 22 and 39, p. 87.
[4] *ACMA Report to Government*, finding 24, p. 87.
[5] *ACMA Report to Government*, finding 26, p. 87.
[6] *ACMA Report to Government*, finding 27, p. 87
[7] *ACMA Report to Government*, finding 28, p. 87
[8] *ACMA Report to Government*, finding 27, p. 87
[9] ACMA (21/03/2022), *Report to government on the adequacy of digital platforms' disinformation and news quality measures*,
https://www.acma.gov.au/report-government-adequacy-digital-platforms-disinformation-and-news-quality-measures (*ACMA Report to Government*).
[10] *EU Strengthened Code of Practice on Disinformation Https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation*
[11] DIGI, *ACPDM, 2022 Review of Australian Code of Practice on Disinformation and Misinformation, Discussion Paper*, June 6 2022

➢ Encouraging greater participation in the code by smaller digital platforms, including by modifying the transparency reporting requirements for services with less than one million active monthly users in Australia.
➢ An updated definition of 'harm' in relation to mis and disinformation, addressing stakeholders' concerns that the threshold of 'serious and imminent' threat of harm was too high; the new threshold is 'serious and credible' threat of harm.
➢ Additional commitments reflecting updates to the strengthened EU Code of Practice in relation to recommender systems, and deterring advertisers from repeatedly placing digital advertisements that propagate mis- and disinformation. There are also updates to further clarify that both sponsored content and paid for advertising are in scope of relevant commitments on demonetising mis and disinformation.
➢ Retaining the pre-existing exclusion of professional news content from being treated as misinformation under the code, and the pre-existing obligation for signatories to address this content when it is being disseminated as disinformation. The review concluded that the ACMA and the professional news media are best placed to address misinformation concerns within their self regulatory and co-regulatory codes.
➢ Requiring greater transparency around the specific products and services that are within scope of the signatories' code commitments, through updates to the code, transparency reporting requirements and the DIGI website.

In addition, the DIGI website was updated in order to improve transparency around signatories' opt-in commitments and the governance arrangements for the code. These improvements are discussed in part 2 below.

Part 2 | **Updates to Transparency reporting process**

In this section of the report, we discuss the updated Best Practice Reporting Guidelines that are included in full in this report in Appendix A. Objective 7 of the ACPDM states: 'Signatories publicise the measures they take to combat Disinformation and Misinformation' and the corresponding outcome states: 'The public can access information about the measures Signatories have taken to combat Disinformation and Misinformation'. Therefore the purpose of the transparency reports is to give the Australian public, the ACMA, and the Australian Government the tools to assess how the signatories are adhering to their commitments under the ACPDM. The goal of the reports is to also help improve understanding of online misinformation and disinformation in Australia over time.

In March 2022, DIGI strengthened the approach to transparency reporting with the introduction of Best Practice Transparency Reporting Guidelines. The Best Practice Transparency Reporting Guidelines were developed by independent expert Hal Crawford, an independent expert in digital news media, appointed under the ACPDM governance arrangement. While they are recommendations and the signatories are not bound to adhere to these recommendations when drafting their reports, they aim to support the outcomes based approach of the ACPDM.

In 2023, DIGI asked Hal Crawford to update the Best Practice Transparency Reporting Guidelines to ensure they align with the revisions made to the ACPDM in December 2022, as detailed earlier in this report. The main changes between the revised 2.0 version guidelines and the version 1.0 guidelines document are:

● The recommended maximum length is now 10,000 words;
● Reference is made in the guidelines to the EU's 2022 Strengthened Code of Practice on Disinformation;
● Reference is made to the need for more trended data, going forward;

- A short section has been added to explain the December 2022 changes to the ACPDM and the revised commitments;
- The reports are to contain a summary of positive progress made in last year's reports; and
- The table with relevant information about signatories commitments under the ACPDM at the commencement of the report has been updated.

All signatories are encouraged to use the Best Practice Transparency Reporting Guidelines when drafting and publishing their fourth annual reports for the period January 2023 to December 2023, noting that the changes to the ACPDM (including the new commitments) came into effect in December 2022. Some signatories elected to use the new format for the 2022 reports where they hold data that is relevant to the commitments made in the revised code[12].

Part 3 | **The 2023 Transparency reports**

2023 Reports published on DIGI website

The transparency reports are a critical aspect of the ACPDM that aims to drive improvements and accountability in the approach of major digital platforms to safeguarding the community from harmful misinformation and disinformation online, through an outcomes based approach that requires annual reporting on signatories' progress[13]. This year's transparency reports have been published on the DIGI website cover the reporting period January 1, 2022 to December 31, 2022 (including the period of the 2022 federal election campaign). The managing director's letter at the beginning of this report highlights some of the notable insights provided by the 2023 transparency reports.

The reports illustrate the volume of misinformation takedowns in Australia, and changes over time. For example:
- **TikTok** has seen improvements in its 'proactive removal' rates of misinformation content – its enforcement of content prior to users even *viewing* it increased from 37.6% in Q1 to 69.8% in Q4 of 2022. In Q4, 89.1% of harmful misinformation was removed before a user *reporting* it.
- **Google** disrupted a Chinese influence campaign known as 'dragonbridge' that perpetuates disinformation on news topics such as China's COVID-19 response to the war in Ukraine. Despite this content receiving relatively low engagement from users, in 2022 Google disrupted over 50,000 instances of 'dragonbridge' activity across its services, and terminated over 100,000 accounts.

The reports detail the nature and effectiveness of interventions to counter misinformation. For example:
- **Meta** has reported on its work to display warnings on content that is found to be false by independent third-party fact checkers; over 2022, it displayed warnings on over 9 million distinct pieces of content in Australia.
- **RedBubble** uses third-party machine learning software to detect users who use bots to create networks of multiple accounts and attempt to upload large amounts of images that may cause public harm.
- **Twitter** details the shifts in its approach after its transition as a company in late October 2022, including the role of its 'community notes' feature that enables Australian users to offer context in

---

[12] The provision of this additional information does not change the date upon which the new commitments under the revised ACPDM come into effect i.e. December 2022.

[13] The first set of reports were filed by eight signatories on May 22, 2021. The second round of annual reports were published on May 30, 2022 and cover signatories activities under the code for the reporting period January 2021 to December 2022#. These reports are available on DIGI's website.

surfacing credible information. If enough contributors with different points of view rate a Tweet as helpful, this is indicated on a Tweet.

The reports illustrate efforts to elevate reputable content and provide users with information to counter misinformation. For example:

- **Apple News** informed and educated readers by sending push notifications for stories sharing important updates about the emergence of COVID-19 Omicron subvariants and other variants.
- **Adobe** released a suite of open-source developer tools based on the C2PA specification, enabling more developers to integrate content provenance across web, desktop, and mobile projects so that consumers can see the origins of the content they are viewing.
- **Microsoft's** search engine Bing applied defensive search interventions for over 45,000 distinct queries impacting over one million impressions in Australia between February 2022 through 31 December 2022.

Independent assessment of the 2023 ACPDM transparency reports

By: Hal Crawford, Crawford Media Consulting

Note on nomenclature: Following last year's precedent, we are calling the reports filed this year the "2023 reports", although they focus on operations within calendar year 2022. The actual reports filed by Signatories this year are in many cases titled "2022 report" for this reason. As required by the guidelines, all reports make clear which period they refer to.

The transparency reports filed this year for the most part continue to incrementally improve on the provision and presentation of information on the efforts of the Signatories to counter mis/disinformation on their platforms. All reports meet the formal requirements of the Code, and many reports have improved in complying with the Transparency Reporting Guidelines.

The key areas for remaining improvement remain the same: consistent metrics reporting, and the provision of trended data. The former sets the foundation for meaningful comparison year-to-year, and the latter provides it. It is understandable in many cases that reported metrics have varied, given changing conditions and the somewhat fluid nature of mis/disinformation. For example, the war in Ukraine was a big event that required a dynamic response from many Signatories.

The situation around trended metrics would be improved if Signatories committed ahead of time to the provision of certain relevant numbers each year. In one case, Twitter, data provided covers only the first half of 2022. Twitter says this is because after the platform's re-invention in late 2022, it is revisiting the way it provides transparency reporting.

It is clear that in the case of all big platforms, the reports represent summations of ongoing activity that is massively complex. Collating the information is a substantial task, and presenting it in a way that balances accuracy and readability is not easy. I am happy that most of the Signatories have struck this balance. They have also heeded the request to include case studies in order to give life to general statements and policies.

Part 4 | **Code Administration**

This section contains an overview of the key activities of DIGI in its role as administrator of the ACPDM.

Overview of complaints received

The complaints facility is designed to provide an added layer of public accountability for signatories to the ACPDM; its goal is to provide a mechanism for Australians to escalate concerns about signatories' commitments to addressing mis-and disinformation, rather than providing an additional avenue for complaints about specific items of online content.

To date, DIGI has received 14 complaints through the complaints portal since it launched in October 2021, of which 9 have been received since the last reporting period. Based on the information provided by complainants to date, complaints have generally been incomplete and/or related to individual items of content on signatories' products or services. None of the complaints have been eligible as potential breaches of the signatories' ACPDM commitments.

DIGI routinely contacts complainants for more information, so as to make assessments as to whether the complaints are in scope for resolution in consultation with our independent complaints sub-committee. DIGI's responses often direct complainants to report content and account issues to the relevant signatory, and where appropriate we refer complainants to relevant Government agencies.
Given DIGI's oversight of the code, we occasionally receive emails of a general nature outside of the complaints portal in relation to content considered to be misinformation. DIGI's replies to such emails usually involve directing the senders to the complaints portal, or suggesting they raise the issue directly with the signatory.

Annual Events

The ACPDM includes a commitment in section 5.29 by signatories to convene an annual event to foster conversations about mis- and disinformation. DIGI convened an inaugural annual event on behalf of signatories at parliament house in May 2021, which featured a panel event discussion of experts and industry about the newly launched Code. This was followed by a virtual annual event in July 2022 that featured a presentation by Jim Reed of Resolve Strategic which provided insights about his findings from his nationally representative survey of Australians' perceptions of misinformation in March 2022. The survey was commissioned by DIGI to help inform the 2022 code review. In 2023, DIGI intends to host an in-person roundtable expert panel discussion about moderating misinformation..

Promotion of Code

Since the launch of the ACPDM, there has been domestic and international interest in the code as a principles based regulatory model. DIGI has engaged with interested Australian stakeholders including consumer groups, academic and online community moderators. In addition, we have shared insights about our experience developing and administering the code and participated virtually in expert roundtable conferences about the management of misinformation in Taiwan, New Zealand, Sri Lanka[14].

---

[14] ACCAN National Conference, Keynote: *Policy approaches to misinformation and disinformatio*n, September 2021 LIRNEAsia *Tackling online misinformation while protecting freedom of expression,* expert roundtable October 2021, Internet Governance Forum, roundtable *Best Practices in Content Moderation* (sponsored by Meta) , December 2021, Aspen Institute New Zealand, *Misinformation Challenges: Systems Thinking and Values based Leadership Seminar*, September 2021, Australian National University, National security college Professional Development Program, *The New Weapons: Propaganda, Misinformation and Fake News,* July 2021, and *Propaganda and InformationWarfare,*

August, 2022 Taiwan Internet Governance conference, September 2022, Australian Community Managers conference, Swarm, Roundtable on mis and disinformation, August 2022,

Since the last annual report, DIGI contributed in an Asian online safety forum, in Singapore[15] and has provided information to parliamentary inquiries that have called for information about mis and disinformation and related topics[16].

DIGI has developed social media advertising plan to grow greater awareness of the code among Australians. These promoted posts across relevant signatory platforms will run at a staged cadence, in order to enable evaluation, and promote awareness of the code's existence, as well as DIGI's complaints function. DIGI also continues to promote key milestones in its governance of the code through media releases and other communications materials.

Governance Committees

The governance arrangements for the ACPDM are set out in Appendix B of this report. In accordance with these arrangements, the Administrative Committee (including the independent committee members met twice and the Steering group met on 3 occasions during the period between 22 May 2021 and May 29 2022.

---

[15] Australian Internet Coalition *Online Safety Forum*, Singapore December 2022.
[16] DIGI submission to House Select Committee, *Inquiry on Social Media and Online Safety* (dated January 12, 2022), DIGI's submission to the *Select Committee on Foreign Interference through Social Media* (dated March 7 2023), DIGI letter to Finance and Public Administration References Committee, *Inquiry concerning the administration of the referendum into an Aboriginal and Torres Strait Islander Voice,* (dated May 17 2023).

Appendix A | **Updated Transparency Reporting Guidelines for The Australian Code on Disinformation and Misinformation**

By: Hal Crawford
Crawford Media Consulting
These guidelines were finalised in May 2023

Prepared for DIGI

Version 2.0, March 2023

This document provides guidelines for Signatories preparing the transparency reports required by the Australian Code of Practice on Disinformation and Misinformation.

The reports themselves fulfill two functions: to inform the public and to provide a framework for the review of activities under the Code.

Two years of reports have already been filed, with the second showing considerable improvement over the first in terms of both functions.

This report urges Signatories to provide:

- Trended data relevant to the Australian market over extended periods
- Consistency in reported metrics year-on-year
- Audience-friendly documents with a minimum of promotional language

In addition, the guidelines list areas for focus and provide a template for the completed reports. The template follows the same form as the 2022 reports, with some additions to reflect the increased scope of the Code from December 2022.

Introduction

The purpose of this document is to set out guidelines for the annual reporting required of Signatories under the Australian Code of Practice on Disinformation and Misinformation ("the Code"). The annual reports are required under Objective 7 of the Code: "Signatories publicise the measures they take to combat Disinformation and Misinformation."[1]

Signatories filed initial annual transparency reports in May 2021, followed a year later by the second series of reports, which covered January-December 2021.

The general purpose of the reports can be inferred from Objective 7 and the Code's administrative requirements, and is twofold:

- To communicate to the general public measures taken by Signatories against mis/disinformation
- To provide a framework for the independent reviewer, DIGI and other stakeholders to audit compliance with the Code

Although these aims are related, they could result in significantly different outputs if one function dominated. A key reminder for signatories is that the documents must be accessible and comprehensible to the general public. The dual purpose of the reports will influence the reporting template recommended in this document.

This is the second iteration of these Guidelines following the introduction of Code. This document responds to the expansion of the Code in December 2022, and removes or revises some of the context provided previously in order to prolong its relevance.

Feedback from past reports

Global regulation of dis/misinformation has developed significantly since work on the Australian Code began, and there have been regulatory developments within Australia that may also affect the operation of the Code in future. For the purposes of this document, we will ignore these external factors and focus on the feedback given to signatories and their responses to it in filing subsequent transparency reports.

The table below indicates the word-count and commitments to outcomes made by each signatory under the Code, as indicated in transparency reports filed in May 2022.

	Wordcount	1	1A	1B	1C	1D	2	3	4	5	6	7
Adobe	3000	✔						✔	✔		✔	✔
Apple	1400	✔	✔		✔		✔		✔		✔	✔
Google	7600	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
Meta	10,200	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
Microsoft	8600	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
Redbubble	2,300	✔	✔	✔	✔		✔	✔			✔	✔
TikTok	4900	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
Twitter	11,000	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔

Good progress was made between the first and second iterations of the transparency reports. Version 1.0 of these guidelines asked that signatories:

- Reduce emphasis on process and policy
- Increase use of trended Australia data
- Adopt a common reporting period
- Use common definitions
- Be explicit with objective/outcome commitments
- Explain reporting metrics
- Provide multi-year metrics reporting
- Increase public accessibility
    - Stick to a word limit (<8000 words)
    - Use breakout case studies
    - Use tables, graphs and other visual elements
- · Reduce promotional tone

Signatories for the most part made improvements in these areas. In particular, the reports fulfilled the request to reduce emphasis on process, to be explicit about committed outcomes, to reduce promotional language and to strive for overall clarity. The word-limit guidance was not universally followed, and improvement in the provision of trended data is still needed. It should be noted that given the signatories diverse business models and technological bases, uniformity in reporting is not possible.

All signatories fulfilled the formal requirements of the transparency reporting regime.

Changes in Code and new obligations

There have been two significant changes to the Code since the last reporting round: the introduction of a requirement to provide information about recommender systems (Outcome 1e) and an addition to Outcome 2, strengthening the reduction of monetisation incentives for dis/misinformation (5:15).

Signatories who commit to these new outcomes will have to include them in their transparency reporting documents. These changes are reflected in additions to the report template given below.

Outcome 1e gives rise to a reporting obligation to show how signatories have "made information available to the end-user" regarding the operation of recommender engines (5.14.A). Signatories who commit to this outcome should also include evidence they have made recommender options available to users (5.14.B).

In terms of Outcome 2, signatories should speak to their efforts to deter advertisers "from repeatedly placing digital advertisements that propagate Disinformation or Misinformation".

Key guidelines

Calendar year reporting

Reports should refer to data from the previous calendar year. The reports filed in May 2023, for example, should relate to the 12 months from 1 January to 31 December 2022. Given that months will have elapsed from the end of that period to the time of compiling the report, it may be acceptable to refer to developments in dis/misinformation in the first half of the current calendar year in passing commentary.

The use of a common reporting period is essential in terms of making comparisons year-on-year and platform-to-platform, and increases the utility of the reports.

Statement of commitments and relevant services/products/platforms

Signatories must state near the beginning of their reports which Code Objectives/Outcomes they have committed to, and which services and products the commitment applies to. This is particularly important for Signatories with big and differentiated portfolios. The omission of a relevant service/product should be noted.

Data in context

In general, there is a need for more trended numerical data in the transparency reports. A minimum of three years of reporting should be supplied with any data, in order to give context. This is a key requirement for understanding.

It may not always be possible to supply three years of data for a given metric. In that case, contextualisation through trended monthly numbers may be appropriate.

Accompanying commentary is vital to explain changes. For example, the incidence of detected dis/misinformation may have increased in a given year because the quantum of dis/misinformation increased, or because a Signatory improved detection. Regardless of the potential misinterpretation of trended data, a transparency reporting regime demands it, and furthermore demands that the same data

be reported in subsequent years. Any addition or omission of data should also be the subject of an explanatory note.

The expectation of the report review process is that data used as internal Key Performance Indicators in the area of dis/misinformation be included in the report unless there is a clear commercial imperative to omit (see more on KPIs below).

Australian data

Reporting under the Code should provide data for the Australian market. Global metrics may also be relevant, but given the Code's national nature the primary concern should be Australian numbers, examples, and context. It is recognised that Australian data may not always be available: if this is the case, the Signatory should explicitly note that this is the case.

Public accessibility

There are other aspects of the initial reports that can be built on to improve communication with a general audience:

·   The emphasis on brevity in the first version of these guidelines was perhaps too severe given the big scope of some Signatories' operations. Limiting to 10,000 words should be possible, however.

·   The use of graphical elements such as bar and line charts helps in communicating numerical information

·   Breakout (separated from main text body) case studies are recommended to illustrate key points and developments

Promotional language

One noticeable aspect of some of the very first reports was a "promotional" tone. It is natural that Signatories seek to portray their efforts and accomplishments in the best possible light. Unfortunately, promotional language undermines the informational content of the reporting, and encourages cynicism towards what are in fact major and important efforts to curb mis/disinformation. We encourage Signatories to avoid promotional writing and to maintain a neutral stance, highlighting problems and successes with equanimity, and thereby increasing the credibility of the reporting.

We appreciate this can be difficult with a public document: a good rule of thumb is to avoid statements and words that would not be found in internal company reporting.

Generic information

Generic information relating to dis/misinformation process and policy that is unchanged from past reports should be condensed or moved to appendices where appropriate. This is to place greater emphasis on novel aspects of the fight against mis/disinformation, and to avoid losing the novel information among material that is the same year-to-year.

The previous iteration of these guidelines found that on average, 84% of the content in the initial 2021 reports was generic information relating to dis/misinformation process and policy.

Bearing in mind the dual purpose of the reports – to communicate to the public and demonstrate compliance with the Code – it is necessary to include some of this information repeatedly. For example, it is a mandatory requirement of reporting that Signatories provide links to reporting mechanisms for dis/misinformation (see below for mandatory links). Signatories may also want to include important information about their approach to tackling dis/misinformation in every report, and there should be an opportunity to do this.

Mandatory links reporting

Signatories' commitments under the Code include simple links to information in order that the independent reviewer may assess compliance. To be explicit regarding these mandatory requirements, they are:

- 1b: Links to user guidelines, policies and procedures relating to mis/disinformation
- 1c: Links to publicly available tools for reporting mis/disinformation
- 5: Links to/evidence of published information that allows users to better distinguish factual information from mis/disinformation

The link requirements are provided as a checklist to ensure simple elements are not omitted. As indicated in the rest of these guidelines, Signatories are expected to elaborate significantly through the identification and provision of relevant data and commentary. Signatories who have not committed to an Outcome/Objective are exempted from the relevant mandatory elements. Note that in addition to providing these links to assist the independent reviewer, it may be helpful for the reviewer to directly query the Signatory on elements of a submitted report.

The issue of KPIs

In the European Union's 2022 Strengthened Code of Practice on Disinformation, great emphasis is put on quantifying the effectiveness of dis/misinformation countermeasures through Service Level Indicators (SLIs) and Qualitative Reporting Elements (QREs) associated with commitments.[2] The idea is that these data may provide a measure of cross-platform comparison. The EU requires that Signatories provide data for the 6-month reporting period on a country-by-country basis.

In practice the first batch of reports filed under the 2022 EU Code indicate that platforms may have had difficulty in compiling or providing these numbers, and that comparison between different organisations is not very informative.

It is likely the most important aspect of the EU reporting regime will be the trends revealed over time within the same organisation. This stands in contrast to the idea of a meaningful system of pan-industry dis/misinformation metrics (Key Performance Indicators, or KPIs).

In keeping with the first iteration of these guidelines, we urge the Australian Code's Signatories to identify and commit to appropriate internal KPIs that are consistently reported on from one year to the next.

Challenges in reporting

The Signatories are diverse businesses and there are big variations in the application of the Objectives/Outcomes. It is not possible to be prescriptive in dictating the data supplied in the

transparency reports, although it is expected that Signatories themselves identify relevant data and supply it in line with the suggestions of this document (i.e., within the Australian market, for the reporting period, and for a minimum of two years prior to that). A particular challenge may arise for the Signatories whose dis/misinformation operations are extensive. We encourage them to focus on changes within the reporting period and their interpretations, responses and initiatives. There are also some Signatories whose activities relevant to the Code cannot be quantified. In this case, Signatories are encouraged to report case studies and such qualitative information as will increase the general understanding of their efforts.

Note on formatting

We recommend Signatories use their own formatting conventions in terms of font, layout and colour in the final PDF document. This will not hinder independent review and may enhance messaging to the general public. As implied in the template below, the reports should follow a common structure, with considerable leeway for different elements like graphs, tables and breakout case studies. This will ensure a degree of uniformity across Signatories and better enable comparison between reporting periods. Where numerical data can be supplied, the preference is to present this at the beginning of sections and to contextualise with commentary. It is important to clearly explain metrics and the rationale behind them.

Report structure

We recommended following the framework below in preparing reports. Content suggestions and constraints are given in brackets. Note the positions of graphs and breakout case studies are given as examples only.

Australian Code of Practice on Disinformation and Misinformation

[Name of Signatory]

Annual Transparency Report

[Reporting period]

Summary

[Discuss in brief the overall features of the reporting period]

[Include analysis of the general environment relevant to dis/misinformation]

[Reiterate the primary elements of your work against dis/misinformation]

[<800 words]

Commitments under the Code

[Use a table to summarise commitments and the platforms they apply to, as below]

1a [paraphrase Outcome 1a]	[platform] [service] [product]
1b [paraphrase Outcome 1b]	[platform] [service] [product]
1c [paraphrase Objective 2]	[platform] [service] [product]
Etc. …	Etc. …

[Include short commentary on omitted objectives/outcomes/platforms/services/products]

Reporting against commitments

Outcome 1a: Reducing harm by adopting scalable measures

[Metrics reported and for what reason]

[Comments on trends observed]

[Any changes in type of content/behaviour targeted]

[Changes to acceptable use policy etc.]

[What measures were successful and how is that reflected in the data?]

[Tables and graphics as appropriate]

[Case studies as appropriate]

CASE STUDY 1

[Illustrates a particular aspect of data trend or impact of changes made]

[Note this is an example location for a case study. If appropriate and available, Signatories should provide several case studies. Such qualitative content is valuable in bringing policy to life.]

Outcome 1b: Inform users about what content is targeted

[What new initiatives in communicating to users what constitutes mis/disinformation?]

[Evidence of user engagement with this content]

[Links to user guidelines, policies and procedures relating to mis/disinformation]

Outcome 1c: Users can easily report offending content

[Any changes in the way users report content for the reporting period]

[Links to publicly available tools for reporting mis/disinformation]

Outcome 1d: Information about reported content available

[What data have you published to users about the amount and quality of dis/misinformation reporting under 1c?]

[Include such data if available]

[Also give links to where the data has been published]

Outcome 1e: Information about recommender systems

[What information have you provided to users about how recommender systems work on your platforms?]

[What options do users have around recommender systems , and how has that been communicated to them?]

[Provide links where possible, or example screenshots if not]

Objective 2: Disrupt advertising and monetisation incentives for disinformation.

 [Explain KPIs as above]

[Quantify progress made against the monetisation of disinformation, graphically if possible]

[Per 5:15, what measures have been taken against advertisers who repeatedly provide ads containing dis/misinformation?]

[Changes to policies and processes implemented to reduce monetisation for targeted content and behaviour]

[Any relevant changes in market conditions]

Objective 3: Work to ensure the integrity and security of services and products delivered by digital platforms.

[Detail of work in the period against inauthentic behaviours that impact product security]

[As above, detail trends and initiatives, and plans in this area]

[This section may contain reference to 1a, given potential overlap in these Objectives – it is acceptable to simply refer to that section if all actions against inauthentic user behaviour are covered there]

Objective 4: Empower consumers to make better informed choices of digital content.

[Detail the ways in which you have helped users distinguish dis/misinformation from quality information]

[What is the uptake or awareness of such "empowerment tools"?]

[In what content categories are they active?]

Objective 5: Improve public awareness of the source of political advertising carried on digital platforms.

[Detail the ways in which you have flagged political advertising and improved the awareness of political sources of advertising]

[Any challenges on the horizon, e.g. Upcoming elections]

CASE STUDY 2

[Illustrates a particular aspect of data trend or impact of changes made]

[Empty box]

Objective 6: Strengthen public understanding of Disinformation and Misinformation through support of strategic research.

[Suggest the use of the table here]

[Name of university/institute/company]	[Overview of research]
…	…
…	…

[Notable success/challenges/changes in the above work]

[Include links]

Objective 7: Signatories will publicise the measures they take to combat Disinformation.

[Aside from this report, what other information about your work against dis/misinformation has been communicated to the public?]

[Quantify engagement with this information if possible]

[Overlaps to some extent with 1d, and if there is complete overlap simply refer to that section]

Concluding remarks

[Unanswered questions and challenges]

[Summary of any new initiatives not already mentioned]

[Evolution of your business's understanding of the problem and how to tackle it]

[Observations on the Code and the process of reporting]

[May include developments between the end of the reporting period and now]

Appendix

[1] "… each Signatory will provide an annual report to DIGI setting out its progress towards achieving the outcomes contained in the Code which will be published on the DIGI website." 7.3, page 17.

[2] Commitment 40, *The Strengthened Code of Practice on Disinformation 2022*

Appendix B | **Governance arrangements for The Australian Code on Disinformation and Misinformation**

In October 2021, DIGI announced the governance arrangements for the ACPDM in order to strengthen the code and its effectiveness. These are summarised here. The code is a novel self regulatory mechanism that aims to drive improvements through increased transparency about how platforms tackle mis and disinformation; DIGI's governance arrangements have been tailored with that aim in mind.

Complaints committee

The Complaints Committee is independent and resolves complaints about possible breaches by signatories of their commitments under the code. DIGI acts as secretary on this committee, but has no vote on decisions in order to avoid conflicts of interest. The committee meets to hear complaints of material code breaches that cannot be resolved by signatories and complainants. The Terms of Reference for the Complaints Sub-committee can be found on the DIGI website[17], and the three independent members of the Complaints Sub-committee are detailed below.

Administration committee

The Administration Sub-Committee brings together the three independent representatives from the Complaints Sub-Committee with signatories of the code. This committee monitors the various actions taken by signatories to meet their obligations under the Code, such as the operation of the complaints facility.

Signatory steering group

As any digital company can adopt the code, not just DIGI's members, this group enables companies that are not members of DIGI to have an equal say in decisions that are made about the code, if they choose. This group serves to separate DIGI's advocacy work on behalf of its members from the code governance functions.

---

[17] DIGI, *Terms of reference for Complaints Facility and Complaints Sub-committee | The Australian Code of Practice on Disinformation and Misinformation*, https://digi.org.au/wp-content/uploads/2021/10/DIGI-TOR-for-Complaints-Facility-and-Complaints-Sub-committee-_-ACPDM-_-FINAL-NE-1.pdf

Independent review of transparency reports

An independent expert fact checks all signatories' transparency reports and provides an attestation of them, in order to incentivise best practice and compliance. The reviewer provides advice to the Administration Sub-committee if it cannot provide an attestation of claims in a transparency report. The attestation process does not involve an evaluation of the quality of the reports or the compliance with the Code, but provides independent confirmation that certain publicly verifiable information is provided in accordance with agreed reporting guidelines. Signatories may also provide an internal contact with whom the reviewer can confidentially verify any internal policies and processes that are not publicly verifiable. The reviewer's role entails:

1. Verifying if each signatory has published and implemented policies and processes that comply with their obligations in sections 5. 8, 5.10, 5.11 and 5.13 that pertain to Objective 1 (Safeguards against Disinformation and Misinformation) and Outcomes 1a, b, c and d of the Code. These sections contain the baselines requirements to implement measures that contribute to reducing the risk of users exposure to Disinformation and Misinformation, explain prohibited behaviors, provide mechanisms to report Disinformation and Misinformation, and provide general information on actions taken in response to reports.
2. Verifying if each signatory has published and implemented policies and processes that comply with their obligations in relation to any optional commitments they have made under the Code.
3. Verifying if the policies and processes mentioned in the transparency report are accessible to Australian users.
4. Verification of 1, 2 and 3 involves checking information provided in the transparency report against public sources.
5. Verifying if each signatory is meeting the ACPDM's commitments regarding the form of the reports including the Best Practice Guidelines.
6. Verification will not involve review of sensitive or proprietary information such as the deployment of technological solutions to detect and remove accounts propagating disinformation.
7. Advising each signatory on a confidential basis if they can attest that the report meets these review requirements, or if there are any gaps.
8. Providing advice to the Administration Sub-committee if they cannot provide an attestation in relation to a signatory's reports, in which case the signatory/ies must either amend and resubmit the reports to the reviewer for further assessment or provide written reasons as to why they dispute the reviewer's assessment.
9. Providing a generalised assessment of the reports, which has been included below in this annual report.

Independent Members of Administration Committee and Complaints Committee

Dr. Anne Kruger

Dr Anne Kruger sits on the Complaints Sub-Committee and the Administration Sub-committee.

Dr Anne Kruger was appointed Associate Director of RMIT FactLab in January 2023 where she brings her expertise in leading the strategic growth of international verification projects. RMIT FactLab is a research hub dedicated to debunking misinformation online and developing critical awareness about its origins and spread. RMIT FactLab is IFCN accredited and also houses RMIT ABC Fact Check and Meta's third-party fact checking.

Anne spent nearly four years with global verification experts First Draft. She now also directs RMIT FactLab's CrossCheck initiative which carries on in APAC the mission of First Draft's pre-emptive collaborative CrossCheck projects as a hub for dozens of external media organisations. First Draft's mission has now moved to Brown University as the newly established Information Futures Lab and RMIT FactLab's CrossCheck is their APAC affiliate.

Anne was co-chief investigator at the University of Technology Sydney's Centre for Media Transition which, together with First Draft, assisted DIGI with the development of its disinformation industry code. Anne was an anchor at CNN Hong Kong during SARS, and later a finance reporter at Bloomberg. As an Associate Professor of Practice, Anne established a verification lab at the University of Hong Kong collaborating with Meedan, taught news literacy at HKU and led

Victoria Rubensohn AM

Victoria Rubensohn AM sits on the Complaints Sub-Committee and the Administration Sub-committee.

Victoria has extensive experience in media and communications regulation in Australia and overseas, and with codes of practice in broadcasting, telecommunications, content classification and advertising. Victoria is currently Consumer Director of Communications Compliance Ltd and Principal of Omni Media. From 2011 to late 2020, Victoria was an Independent Reviewer for Ad Standards Australia; From 2015 to 2019, she was a Consumer Member of the Code Authority of ADMA; and from 2009 to 2015, she was Convenor of the Classification Review Board. Victoria has chaired federal government policy review committees on copyright convergence and digital radio. From 1994 to 2009, Victoria chaired the Telephone Information Services Standards Council regulating value-added telecommunications services. Victoria is a Director of the Australian Communications Consumer Action Network (ACCAN), the Centre for Inclusive Design, chair of the Communications Law Centre Ltd and is an Advisory Board Member of the Centre For Media Transition at UTS. She also serves on the advisory committee of the International Institute of Communications Australian Chapter, and was formerly President of the Communications and Media Law Association and Chair of the National Film and Sound Archive.

media literacy projects with UNESCO throughout APAC. She previously held senior editorial and presenter positions with ABC Australia. Anne has a PhD in social media verification education.

Christopher Zinn

Christopher Zinn sits on the Complaints Sub-Committee and the Administration Sub-committee.

Christopher has led various successful and disruptive campaigns to help consumers make better decisions in complex markets such as energy, private health insurance and financial services. Christopher heads the www.determinedconsumer.com initiative, is the CEO of the Private Health Insurance Intermediaries Association, sits on the statutory authority reforming the funeral industry, and is on a self-regulatory code committee for the charitable sector. He was also director of communications and campaigns for consumer group CHOICE and has been a reporter and producer for TV, radio and newspapers both in Australia and overseas including the ABC, the Daily Telegraph, Channel Nine, and the UK Guardian.

Hal Crawford

Hal Crawford is the independent reviewer of the 2021 transparency reports, and author of the Best Practice Guidelines.

Hal is a news executive and consultant with 25 years' experience in digital and broadcast media, and currently runs Crawford Media Consulting. He was recently the News Director at one of New Zealand's biggest newsrooms, Newshub, managing news and current affairs on TV, online and radio platforms. Hal was previously Editor-in-chief at ninemsn, and a non-executive director of youth content publisher Pedestrian. Hal is currently chair of the Advisory Board for the Centre for Media Transition at UTS, and has made other contributions to industry and regulatory frameworks for broadcast and digital media. Hal was the co-founder of a social media news sharing project, Sharewars, which gathered and analysed content data from the world's biggest news publishers. He co-authored the book *All Your Friends Like This*, published by HarperCollins in 2015, about social media and news.

Complaints portal

A key component of the governance arrangements is the public complaints portal, that is available on DIGI's website[18]. The operation of the portal is detailed publicly in the complaints facility terms of reference[19], which explains the processes for how complaints are resolved. The resolution measures have

---

[18] DIGI, *Complaints*, https://digi.org.au/disinformation-code/compaints/
[19] DIGI, *Terms of reference for Complaints Facility and Complaints Sub-committee | The Australian Code of Practice on Disinformation and Misinformation*,

been designed to provide incentives for signatories to address breaches of the code, which is considered a better outcome than more punitive resolution measures.

When a complaint is made through the portal, DIGI assesses its eligibility and escalates the complaint according to a standardised internal process that is overseen and approved by the complaints sub-committee. The complaints form enables members of the public to make complaints where they believe a signatory has breached a code commitment. This approach is consistent with the recommendations of the final report from the ACCC Digital Platforms Inquiry, which recommended an approach to complaints that centres on code breaches through a focus on 'assessing the response of the digital platforms to complaints against the terms of the code'[20].

Signatories to the ACPDM also commit to providing an avenue for the public to make complaints about instances of mis- and disinformation on their platforms. DIGI does not accept complaints about individual items of content on signatories' products or services, and encourages members of the Australian public to report misinformation or materials that violate specific platform policies directly to the code signatories via their reporting mechanisms.

✅ Example of **eligible** complaint	☒ Example of **ineligible** complaint
A failure to implement and publish policies and/or reporting that will enable users to report the types of behaviours and content that violates their policies under section 5.10 of the Code.	A determination by a signatory that specific items of content or categories of content is or is not disinformation or misinformation, or a decision to remove an individual's account. Those complaints will be handled by the signatories under the policies and procedures for reporting issues they are committed to implement under section 5.11 of the code.

---

https://digi.org.au/wp-content/uploads/2021/10/DIGI-TOR-for-Complaints-Facility-and-Complaints-Sub-committee-_-ACPDM-_-FINAL-NE-1.pdf

[20] ACCC (2019), *Digital Platforms Inquiry Final Report*, https://www.accc.gov.au/system/files/Digital%20platforms%20inquiry%20-%20final%20report.pdf, p. 371