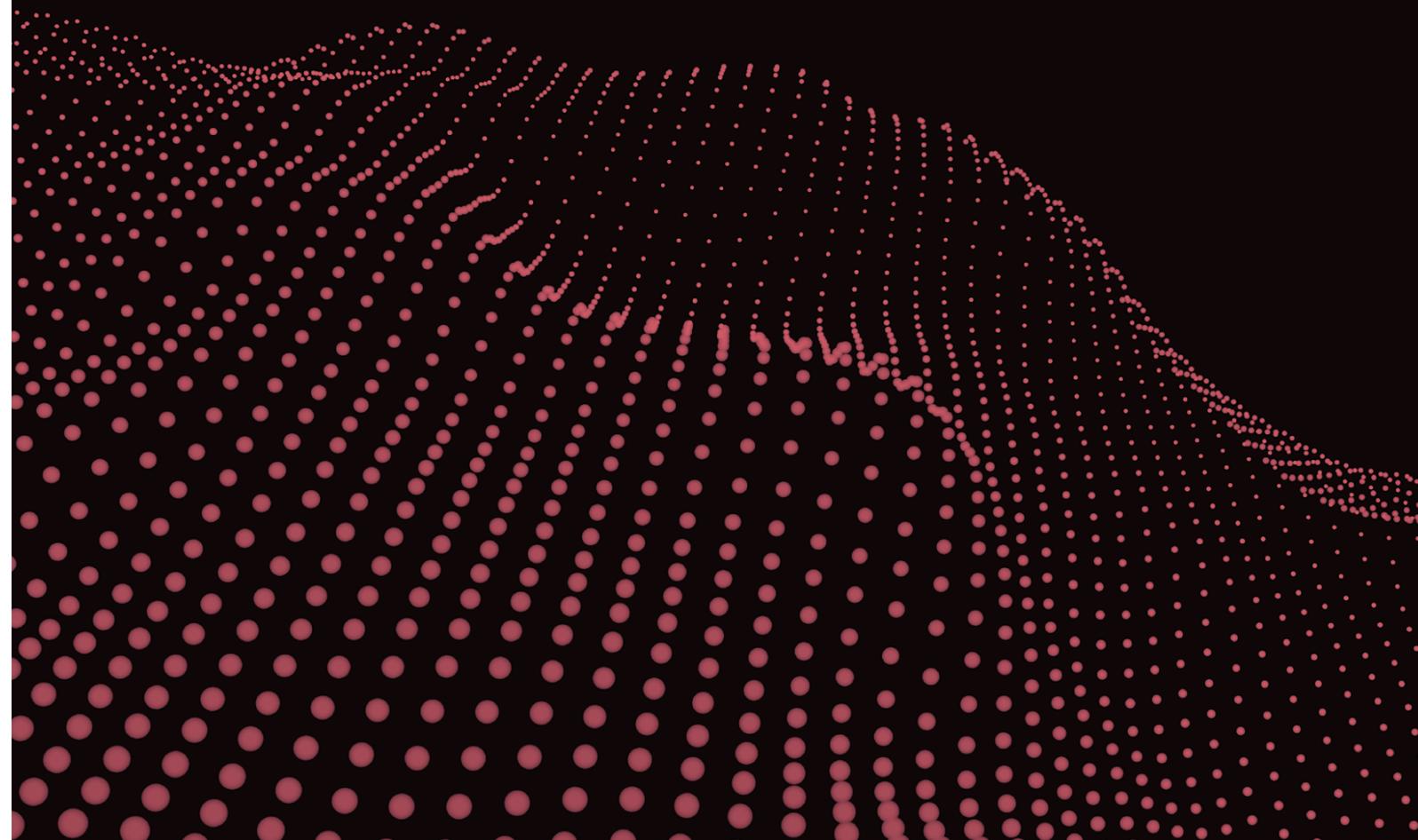




Australian Code of Practice on Disinformation and Misinformation

An industry code of practice developed by the Digital Industry
Group Inc. (DIGI).

Updated December 22, 2022



1. Preamble

- 1.1. *Background:* The Australian Code of Practice on Disinformation and Misinformation (The Code) has been developed by the Digital Industry Group Inc. (DIGI), a non-profit industry association that advocates for the interests of the digital industry in Australia. The Code was developed in response to Government policy as set out in *Regulating in the Digital Age: Government Response and Implementation Roadmap* for the Digital Platforms Inquiry, where Government asked the major digital platforms to develop a voluntary code of conduct outlining what the platforms will do to address concerns regarding disinformation and credibility signalling for news content. The Code also takes into account guidance provided by the Australian Communications and Media Authority set out in *Misinformation and News Quality on Digital Platforms in Australia: A Position Paper to Guide Code Development*. In December 2022, the Code was revised following a review which took into account feedback provided via a public consultation process.
- 1.2. *Subject matter:* Disinformation and misinformation are aspects of a wider, multifaceted social problem which involves a range of offline and online behaviours which propagate information that threatens to undermine established democratic processes or public goods such as public health. Concepts such as “disinformation”, “misinformation”, and “fake news” mean different things to different people and can become politically charged when they are used by people to attack others who hold different opinions on value-laden political issues on which reasonable people may disagree. The understanding and effects of these concepts varies amongst individuals and is also under-researched.
- 1.3. *Role of digital platforms:* The digital platforms who have signed this Code recognise their role as important actors within the Australian information ecosystem and have already implemented a range of measures to tackle the propagation of disinformation and misinformation amongst users of their services and products. This Code is designed to express the minimum commitments made by Signatories to address the propagation of Disinformation and Misinformation (as defined in this Code) via digital platforms.
- 1.4. *Signatories:* This Code may be signed by any digital platform who provides products and services to Australian users and has identified a risk that end-users may propagate disinformation and misinformation online and/or can contribute to reducing the propagation of disinformation or misinformation online.

- 1.5. *Minimum Commitments:* All Signatories commit to meet the commitments outlined in section 5.2 including the core objective of providing appropriate safeguards against Harms that may be caused by Disinformation and Misinformation.
- 1.6. *Opt-in:* The digital industry is highly innovative and diverse, and digital platforms operate vastly different businesses which offer a wide and constantly evolving variety of services and products. As a result, the measures taken by digital platforms to address Disinformation and Misinformation in the context of their respective businesses may vary over time. For example, measures which are taken by a user-generated content platform may differ from those taken by a search engine. To accommodate the need of the Signatories to choose those measures which are most suitable to address instances of Disinformation and Misinformation in relation to different services and products provided by digital platforms, this Code provides Signatories the ability to opt into a range of measures and objectives, additional to the minimum commitments outlined in 1.4.
- 1.7. *Proportionality:* The types of user behaviours, content and Harms that this Code seeks to address will vary greatly in incidence and impact amongst the diverse range of services and products offered by different digital platforms. Accordingly, the commitments made by Signatories to the Code are intended to enable them to take actions which are proportional responses to their commitments under the Code. Section 6 provides further guidance on the contextual factors that Signatories may take into account in this regard.
- 1.8. *Signatories may take a more expansive approach:* Signatories may in their discretion implement policies and processes that contain measures to combat Misinformation and Disinformation on a wider range of content or products and services than is within the scope of this Code.

- 1.9. *Need for collaboration and cooperation among all relevant stakeholders:* While this Code is intended to apply to digital platforms, the Signatories recognise and emphasise that a range of relevant stakeholders have roles and responsibilities in dealing with Disinformation and Misinformation including public authorities, academia, civil society, influencers, and news organisations. Tackling Disinformation and Misinformation effectively will require concerted effort and collaboration by and among these various stakeholder groups, and not only digital platforms. The Signatories welcome ongoing dialogue with stakeholders about what works well, and what does not.
- 1.10. *Best Practice Guidance:* The Signatories encourage other participants in the information ecosystem such as other digital services to use this Code as a guide to best practice in developing their own response to the evolving challenges of Disinformation and Misinformation.

2. Guiding Principles

- 2.1. *Protection of freedom of expression:* Digital platforms provide a vital avenue for the open exchange of opinion, speech, information, research and debate and conversation as well as creative and other expression across the Australian community. Signatories should not be compelled by Governments or other parties to remove content solely on the basis of its alleged falsity if the content would not otherwise be unlawful. Given its subject matter, the Code gives special attention to international human rights as articulated within the *Universal Declaration on Human Rights*, including but not limited to freedom of speech. Signatories are encouraged to, in developing proportionate responses to Disinformation and Misinformation to be cognisant of the need to protect these rights.
- 2.2. *Protection of user privacy:* Digital platforms value their users' privacy. Any actions taken by digital platforms to address the propagation of Disinformation and Misinformation should not contravene commitments they have made to respect the privacy of Australian users, including in terms and conditions, published policies and voluntary codes of conduct as well as by applicable laws. This includes respect for users' expectations of privacy when using digital platforms and in private digital communications. Additionally, any access to data for research purposes must protect user privacy.

- 2.3. *Policies and processes concerning advertising placements:* Digital platforms recognise the importance of having policies and processes in place with respect to advertisement placements on their services and products to reduce revenues that may reach the propagators of Disinformation and Misinformation.
- 2.4. *Empowering users:* Digital platforms should empower users to make informed choices about digital media content that purports to be a source of authoritative current news or of factual information.
- 2.5. *Integrity and security of services and products:* Digital platforms should communicate on the effectiveness of efforts to ensure the integrity and security of their services and products by taking steps to prohibit, detect and take action against inauthentic accounts on their services and products whose purpose is to propagate Disinformation.
- 2.6. *Supporting independent researchers:* Digital platforms recognise the importance of industry support for research efforts by independent experts including academics that can inform on trends and effective means to counter Disinformation and Misinformation. The Code provides various options for digital platforms to participate in independent research initiatives.
- 2.7. *Without prejudice commitments:* This Code is without prejudice to other initiatives aimed at tackling Disinformation and Misinformation by digital platforms.

3. Glossary

This glossary provides information on some of the key terms used in this Code.

- 3.1. *Digital Content* is content distributed online on a platform owned and operated by a Signatory to this Code that is targeted at Australian users and includes content that has been artificially produced, manipulated or modified by automated means such as through the use of an artificial intelligence algorithm.
- 3.2. The aspect of *Disinformation* that this Code focuses on is:
 - A. Digital Content that is verifiably false or misleading or deceptive;
 - B. is propagated amongst users of digital platforms via Inauthentic Behaviours; and

- C. the dissemination of which is reasonably likely to cause Harm.
- 3.3. *Enterprise Services* is software and services including cloud storage and content delivery services which are designed for the use of a specific organisation.
- 3.4. *Harm* means harms which pose a credible and serious threat to:
- A. democratic political and policymaking processes such as voter fraud, voter interference, voting misinformation; or
 - B. public goods such as the protection of citizens' health, protection of marginalised or vulnerable groups, public safety and security or the environment.
- Note: Harm which poses a credible and serious threat excludes harm that cannot be reasonably foreseen.
- 3.5. *Inauthentic Behaviour* includes spam and other forms of deceptive, manipulative or bulk, aggressive behaviours (which may be perpetrated via automated systems) and includes behaviours which are intended to artificially influence users' online conversations and/or to encourage users of digital platforms to propagate Digital Content.
- 3.6. *Misinformation* means:
- A. Digital Content (often legal) that is verifiably false or misleading or deceptive;
 - B. is propagated by users of digital platforms; and
 - C. the dissemination of which is reasonably likely (but may not be clearly intended to) cause Harm.
- 3.7. A news source is a journalistic producer of news that has editorial independence from the subjects of its news coverage and is:
- A. subject to the rules of the Australian Press Council Standards of Practice or the Independent Media Council Code of Conduct; or
 - B. subject to the rules of the Commercial Television Industry Code of Practice, the Commercial Radio Code of Practice or the Subscription Broadcast Television Codes of Practice; or
 - C. subject to the rules of a code of practice other regulatory instrument that specifies standard of editorial practice in another country; or

- D. is subject to internal editorial standards that relate to the provision of quality journalism; or
 - E. provides a publicly accessible mechanism for making requests for corrections or complaints about the quality of its news coverage.
- 3.8. Digital advertising services means paid for digital advertising services where the placement of the advertisement is sold directly by Signatories to advertisers
- 3.9. *Political Advertising* means paid for advertisements:
- A. made by, on behalf of a political party; or
 - B. that advocate for the outcome of an election, referendum or other Federal, State or Territory wide political process (such as a postal vote) supervised by an electoral management body of the Commonwealth or State and Territory.
 - C. are regulated as political advertising under Australian law.
- 3.10. *Professional news* is online material produced by a news source that reports, investigates, or provides critical analysis of:
- A. issues or events that are relevant in engaging end-users in public debate and in informing democratic decision-making; or
 - B. current issues or events of public significance to end-users at a local, regional or national level.
- 3.11. *Recommender system* means a fully or partially automated system used by an online platform to suggest or prioritise in its online interface specific items of Digital Content to recipients of the service, including as a result of a search initiated by the recipient of the service or otherwise determining the relative order or prominence of the items of Digital Content displayed.

- 3.12. *Search Engines* consist of software systems designed to collect and rank information on the World Wide Web in response to user queries. Search Engines automate their systems in two ways. First, they use software known as “web crawler,” “bots” or “spiders” to discover publicly available webpages and automatically index and collect information from and about these webpages and internet sites. Second, they use ranking systems to return results in a set of links to websites. These ranking systems are made up of a series of algorithms that are ranked based on many factors such as relevance and usability of pages, expertise of sources, and more. The weight applied to each factor may depend on the nature of the query. “Search Engine” excludes downstream entities that offer search functions on their own platforms, for which the results are powered by third-party search engines, as these downstream entities have no legal or operational control of the search results, the index from which they are generated nor the ranking order in which they are provided.
- 3.13. *Sponsored content* is a paid arrangement between a social media service and an account-holder under which the social media service promotes content posted on the service beyond the account holder’s list of followers in exchange for payment but excludes paid for advertising, for example, paid advertising on search engines.

4. Scope, application and commencement of this Code

- 4.1. *Scope*: Recognising that the the incidence and impact of the risks posed by Disinformation or Misinformation will vary greatly, amongst the diverse range of services and products provided by Digital platforms in Australia, it is expected that the commitments under this Code will apply primarily to services and products that disseminate Digital Content to end users in Australia where Signatories consider the risk of disseminating Disinformation and Misinformation is material and/or where Signatories consider they can make a the material contribution to reducing the impact of Harms arising from Disinformation and Misinformation. The Code may, therefore, be signed by a broad range of Signatories and a range of products and services, and is not limited to specific types of Digital Content. For example, products services in scope may include those that:
- A. disseminate user-generated Digital Content (including shared content); and/or

- B. disseminate Digital Content that is returned and ranked by Search Engines in response to user queries;
- C. provide technological solutions that aim to assist digital platforms and /or end-users combat Disinformation and Misinformation;
- D. offer sponsored content and/or digital advertising services; and
- E. aggregate and disseminate news and other types of journalistic content from a variety of different sources.

Note for D: political advertising is excluded from the scope of Misinformation but may fall within the definition of Disinformation if propagated by Inauthentic Behaviours

Note for E: professional news content disseminated by a news aggregation service is excluded from the definition of Misinformation but may fall within the definition of Disinformation if propagated by Inauthentic Behaviours.

4.2. *Excluded services and products:* The following are not services and products subject to this Code:

- A. private messaging services including those provided via software applications;
- B. email services including those provided via software applications;
- C. Enterprise Services.

4.3. The list of services and products that are within the scope of the Code or which are excluded is not intended to be exhaustive as new services and products are likely to emerge, some of which will be relevant to the Code and some of which will be irrelevant Signatories will identify the relevant products and services covered by this Code in accordance with section 7.1, 7.3 and 7.4 of the Code.

4.4. *Content that is not Misinformation:* The following content is not Misinformation under this Code:

- A. content produced in good faith for entertainment (including satire and parody) or for educational purposes;
- B. content that is authorised by an Australian State or Federal Government;
- C. subject to sections 5.23 to 5.25, Political Advertising or content authorised by a political party registered under Australian law; and
- D. professional news content.

Content that falls within A to D of this section 4.4 may fall within the definition of Disinformation if propagated by Inauthentic behaviors.

- 4.5. Signatories may, in their discretion, implement policies and procedures which govern the dissemination by users on their platforms of the types of content excluded from the provisions of the Code concerning Misinformation under section 4.4, where Signatories determine such content is reasonably likely to cause Harm.
- 4.6. *Application of existing laws:* There are a range of existing laws or regulatory arrangements (such as the *Online Safety Act 2021 (Cth)*) as well as prohibitions or restrictions concerning matters as diverse as tobacco, therapeutic goods, online gambling, election advertising, and defamation that may overlap with some of the matters covered by the Code. To the extent of any conflict with this Code, those laws and regulations will have primacy.
- 4.7. *Application:* The commitments made by each Signatory apply to it, in respect of the commitments it adopts, in respect of the products and services it nominates, from the date that it opts into those commitments.
- 4.8. *Commencement:* This Code commenced on 22 February 2021. Revisions to the Code were made on October 11, 2021 (to reflect new governance arrangements) and on December 20, 2022 (to reflect the outcome of the Code’s first review).

5. Objectives and Measures

- 5.1. *General:* This section incorporates a range of measures aimed at achieving the Code’s objectives and outcomes which are informed by the purpose and guiding principles set out in section 2 above.
- 5.2. *Signatories Commitments.* All Signatories commit to the core Objective 1 of this Code so as to contribute to reducing the risk of Harms that may arise from the propagation of Disinformation and Misinformation on digital platforms as set out in Outcome 1a and will provide transparency reports as set out in section 7. Not all objectives and outcomes will be applicable to all Signatories who may adopt one or more of the measures set out in this section 5 in a manner that is relevant and proportionate to their different services and products, in accordance with the guidance in section 6. Signatories recognise that measures implemented under the Code may also evolve to reflect changes in their services and products, technological developments and the information environment.

- 5.3. *Opt-in*: Section 7.1 below outlines how Signatories will elect to opt into the commitments.
- 5.4. *Terminology of measures*: In implementing measures under the Code, Signatories recognise that actions taken aimed at achieving any outcome including the implementation of policies and processes may use terminology other than “Disinformation” and “Misinformation” and may, for example, refer to or a range of prohibited user behaviours or conduct such as making false or misleading representations about the user’s identity, origin or intentions and/or a range of prohibited content such as misleading, deceptive, dangerous or harmful content.
- 5.5. *Plain language*: Where Signatories commit to publishing their policies, procedures and any relevant community guidelines or additional information on their actions to combat Disinformation and Misinformation, they will use reasonable commercial efforts to do so in plain language and in an accessible, user-friendly format.
- 5.6. *Restrictions on lawful content or users’ access*: In seeking to comply with the requirements of this Code, Signatories are not required to (although they may elect to) take measures that require them to delete or prevent access to otherwise lawful content solely on the basis that it is or may be misleading or deceptive or false. Nor will Signatories be required to signal the veracity of content uploaded and shared by their users.
- 5.7. *Need for transparency to be balanced against disclosure risks*: Signatories recognise that in implementing commitments to promote the public transparency of measures taken under this Code there is a need to balance the need to be open about those measures with the risk that the release of certain information may result in an increase in behaviours that propagate Disinformation and Misinformation, or which increase its virality.

Objective 1: Provide safeguards against Harms that may arise from Disinformation and Misinformation.

Outcome 1a: Signatories contribute to reducing the risk of Harms that may arise from the propagation of Disinformation and Misinformation on digital platforms by adopting a range of scalable measures.

- 5.8. Signatories will develop and implement measures which aim to reduce the propagation of and potential exposure of users of digital platforms to Disinformation and Misinformation.
- 5.9. Measures implemented under 5.8, may include, by way of example rather than limitation:
- A. policies and processes that require human review of user behaviours or content that is available on digital platforms (including review processes that are conducted in partnership with fact-checking organisations);
 - B. labelling false content or providing trust indicators of content to users;
 - C. demoting the ranking of content that may expose users to Disinformation and Misinformation;
 - D. removal of content which is propagated by Inauthentic Behaviours;
 - E. providing transparency about actions taken to address Disinformation and Misinformation to the public and/or users as appropriate;
 - F. suspension or disabling of accounts of users which engage in Inauthentic Behaviours;
 - G. the provision or use of technologies to identify and reduce Inauthentic Behaviours that can expose users to Disinformation such as algorithmic review of content and/or user accounts;
 - H. the provision or use of technologies which assist digital platforms or their users to check authenticity or accuracy or to identify the provenance or source of digital content;
 - I. exposing metadata to users about the source of content;
 - J. enforcing published editorial policies and content standards;

- K. prioritising credible and trusted news sources that are subject to a published editorial code (noting that some Signatories may choose to remove or reduce the ranking of news content which violates their policies in accordance with section 4.5);
- L. partnering and/or providing funding for fact checkers to review Digital Content; and
- M. providing users with tools that enable them to exclude their access to certain types of Digital Content.

Outcome 1b: Users will be informed about the types of behaviours and types of content that will be prohibited and/or managed by Signatories under this Code.

5.10. Signatories will implement and publish policies and procedures and appropriate guidelines or information relating to the prohibition and/or management of user behaviours and/or content that may propagate Disinformation and/or Misinformation via their services or products.

Outcome 1c: Users can report content or behaviours to Signatories that violate their policies under section 5.10 through publicly available and accessible reporting tools.

5.11. Signatories will implement and publish policies, procedures and appropriate guidelines that will enable users to report the types of behaviours and content that violates their policies under section 5.10.

5.12. In implementing the commitment in section 5.11, Signatories recognise that the terms Disinformation and Misinformation may be unfamiliar to users and thus policies and procedures aimed at achieving this outcome may specify how users may report a range of impermissible content and behaviours on digital platforms.

Outcome 1d: Users will be able to access general information about Signatories' actions in response to reports made under 5.11.

5.13. Signatories will implement and publish policies, procedures and/or aggregated reports (including summaries of user reports made under 5.11) regarding the detection and removal of content that violates platform policies, including but not necessarily limited to content on their platforms that qualifies as Misinformation and/or Disinformation.

Outcome 1e: Users will be able to access general information about Signatories' use of recommender systems and have options relating to content suggested by recommender systems.

5.14. Signatories that provide services (other than search engines) whose primary purpose is to disseminate information to the public and which use recommender systems, commit to :

- A. make information available to end-users about how they work to prioritise information that end-users may access on these services; and
- B. provide end-users with options that relate to content suggested by recommender systems that are appropriate to the service.

Note: for example, the comments section provided under news stories published by an online newspaper would be ancillary to the main service represented by the publication of news under the editorial responsibility of the publisher and therefore not subject to this commitment.

Objective 2: Disrupt advertising and monetisation incentives for Disinformation and Misinformation.

Outcome 2: Advertising and/or monetisation incentives for Disinformation and Misinformation are reduced.

5.15. Signatories that offer digital advertising services will use commercially reasonable efforts to deter advertisers from repeatedly placing digital advertisements that propagate Disinformation or Misinformation.

5.16. Signatories will implement policies and processes that aim to disrupt advertising and/or monetisation incentives for Disinformation or Misinformation.

- 5.17. Policies and processes implemented under 5.16 may for example, include:
- A. promotion and/or inclusion of the use of brand safety and verification tools;
 - B. enabling engagement with third party verification companies;
 - C. assisting and/or allowing advertisers to assess media buying strategies and online reputational risks;
 - D. providing advertisers with necessary access to client-specific accounts to help enable them to monitor the placement of advertisements and make choices regarding where advertisements are placed; and /or
 - E. restricting the availability of advertising services and paid placements on accounts and websites that propagate Disinformation or Misinformation.
- 5.18. Signatories recognise that all parties involved in the buying and selling of online advertising and the provision of advertising-related services need to work together to improve transparency across the online advertising ecosystem and thereby to effectively scrutinise, control and limit the placement of advertising on accounts and websites that propagate Disinformation.

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

Outcome 3: The risk that Inauthentic User Behaviours undermine the integrity and security of services and products is reduced.

- 5.19. Signatories commit to take measures that prohibit or manage the types of user behaviours that are designed to undermine the integrity and security of their services and products, for example, the use of fake accounts or automated bots that are designed to propagate Disinformation.
- 5.20. To allow for the expectations of some users and digital platforms about the protection of privacy, measures developed and implemented in accordance with this commitment should not preclude the creation of pseudonymous and anonymous accounts.

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

Outcome 4: Users are enabled to make more informed choices about the source of news and factual content accessed via digital platforms and are better equipped to identify Misinformation.

- 5.21. Signatories will implement measures to enable users to make informed choices about Digital Content and to access alternative sources of information.
- 5.22. Measures developed and implemented in accordance with the commitment in 5.21 may include, for example:
 - A. the use of technological means to prioritise or rank Digital Content to enable users to easily find diverse perspectives on matters of public interest;
 - B. aggregation or promotion of news content subject to an independent editorial code and complaints scheme;
 - C. the provision or use of technologies which signal the credibility of news sources, or which assist digital platforms or their users to check the authenticity or accuracy of online news content, or to identify its provenance or source;
 - D. the promotion of digital literacy interventions, informed by evidence or expert analysis; and/or
 - E. the provision of financial support and/or sustainable partnerships with fact-checking organisations.

Objective 5: Improve public awareness of the source of Political Advertising carried on digital platforms.

Outcome 5: Users are better informed about the source of Political Advertising.

- 5.23. While Political Advertising is not Misinformation for the purposes of the Code, Signatories will develop and implement policies that provide users with greater transparency about the source of Political Advertising carried on digital platforms.
- 5.24. Measures developed and implemented in accordance with the commitment in 5.23 may include requirements that advertisers identify and/or verify the source of Political Advertising carried on digital platforms; policies which prohibit advertising that misrepresents, deceives, or conceals material information about the advertiser or the origin of the advertisement; the provision of tools which enable users to understand whether a political ad has been targeted to them; and policies which require that Political Advertisements which appear in a medium containing news or editorial content are presented in such a way as to be readily recognisable as a paid-for communication.
- 5.25. Signatories may also, as a matter of policy, choose not to target advertisements based on the inferred political affiliations of a user or choose to define and implement commitments concerning a broader scope of political advertising including advertising that advocates for a political outcome on social issues of public concern.

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

Outcome 6: Signatories support the efforts of independent researchers to improve public understanding of Disinformation and Misinformation.

- 5.26. Signatories commit to support and encourage good faith independent efforts to research Disinformation and Misinformation both online and offline. Good faith research includes research that is conducted in accordance with the ethics policies of an accredited Australian University, provided such policies require that data collected by the researcher is used solely for research purposes and is stored securely on a university IT system, or any research which is conducted in accordance with the prior written agreement of the digital platform.
- 5.27. Measures taken to implement 5.26 may include, for example, cooperation with relevant initiatives taken by independent fact checking bodies. Other measures may include funding for research and/or sharing datasets, undertaking joint research, or otherwise partnering with academics and civil society organisations.
- 5.28. Signatories commit not to prohibit or discourage good faith research, as described in 5.26, into Disinformation or Misinformation on their platform.
- 5.29. Relevant Signatories commit to convene an annual event to foster discussions regarding Disinformation and Misinformation within academia and Civil Society.

Note: The annual event may be conducted via any format (online, offline or a combination) and may cover any topic that is relevant to the Code or its subject matter. The event may, for example, be convened to present or discuss a research initiative of Signatories or the Code administrator.

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

Outcome 7: The public can access information about the measures Signatories have taken to combat Disinformation and Misinformation.

- 5.30. All Signatories will make and publish a transparency report information in section 7.

- 5.31. In addition, Signatories will publish additional information detailing their progress in relation to Objective 1 and any additional commitments they have made under this Code.
- 5.32. Signatories may fulfill their commitment in section 5.31 by providing additional reports and/or public updates on areas such as content removals, open data initiatives, research reports, media announcements, user data requests and business transparency reports. Examples of such information could include, by way of example rather than limitation, blog posts, white papers, in-product notifications, transparency reports, help centres, or other websites.

6. Guidance on platform-specific measures

- 6.1. *Proportionality of measures under Code:* The measures taken by Signatories pursuant to this Code will be proportionate and relevant to their specific context including the Harm posed by instances of Disinformation and Misinformation. Signatories may take into consideration a variety of factors in assessing the appropriateness of measures including:
- A. the actors which are engaged in propagating Disinformation and Misinformation;
 - B. the nature of the behaviour of users propagating Disinformation and Misinformation, for example, whether the behaviour is automated and intentional and/or maliciously motivated and the extent to which it is coordinated, persistent and at scale;
 - C. the type of Product or Service via which the content is distributed and whether it has network effects that result in content being widely and rapidly shared amongst users of the platform;
 - D. whether the platform may receive a commercial benefit from the propagation of the content (for example, whether the content is sponsored content);
 - E. the extent to which it is reasonably possible to verify the falsity of relevant Digital Content via an authoritative or credible source;
 - F. the proximity and severity of the Harm that is reasonably likely to result from the propagation of the content;
 - G. the nature of the online community using the digital platform;

- H. the size and nature of the digital platform’s business and the resources available to it;
- I. the need to protect freedom of expression in balance with other human rights; and
- J. the need to protect user privacy.

7. Code administration

- 7.1. *Opt-in:* In recognition of the variation in business models and product offerings of Digital platforms, this Code is designed to allow a range of businesses to make commitments by way of opt-in arrangements. Within three months of signing the Code, Signatories will nominate the provisions to which they commit using the Opt-in Nominations Form in Appendix 1. A Signatory is not bound to comply with commitments it has not nominated.

Each Signatory will annually re-assess the extent the provisions of the Code are relevant to their products and services (including whether any new products and services should be subject to the Code) and update and notify DIGI of any updates to the opt-in form. DIGI will publish updates to the Opt-in Nominations Form on the DIGI website.

- 7.2. *Withdrawal from Code:* A Signatory may withdraw from the Code or a particular commitment under the Code by notifying DIGI.
- 7.3. *Annual Reporting for Signatories that are large Digital Platforms:* In addition to the Opt-in Nomination Form under section 7.1, each Signatory that has a service subject to the Code with a user base in excess of one million monthly active Australian end- users 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. Each Signatory’s annual report will list its product and services that are subject to the Code including any additional products and services that have been assessed to be subject to the Code during the period covered by the report.

Reporting will be based on the form of the template report provided in Appendix 2 and will be submitted by May 30 each year, covering the period for the previous calendar year.

7.4. *Reporting for Signatories that are smaller Digital platforms*

In addition to the Opt-in Nomination Form under section 7.1, each Signatory that provides a service in Australia that has a user base of less than one million monthly active Australian end- users will provide an initial report to DIGI setting out its progress towards achieving the outcomes contained in the Code which will be published on the DIGI website. This report may be based upon (but need not comply) with the template in Appendix 2 and will be submitted within twelve months of the Signatory entering the Code, and should be updated by the Signatory annually if there are any material changes to the report, such as the addition of new products and services that are subject to commitments under the Code; or changes to the range of measures, policies, procedures and/or guidelines that would materially impact the commitments under section 5.8, 5,10 or 5.11 of the Code .

7.5. *Complaints:* Signatories have established a complaints facility for resolving complaints by the public about possible breaches of Signatories’ commitments under the Code. The public can access the facility via a complaints portal on DIGI’s website. An independent Complaints Committee resolves complaints concerning Signatories compliance with the Code in accordance with Terms of Reference, which are published on the DIGI website, together with information about the operation of the complaints facility and the governance of the Code. The complaints facility does not accept complaints about individual items of content on Signatories’ products or services, which should be directed to the relevant Signatory via their reporting mechanisms.

7.6. As part of this process, Signatories will also consider how they can leverage current arrangements with government and relevant regulatory agencies to identify and address instances of Inauthentic Behaviours that propagate Disinformation and are the subject of measures addressed by this Code.

7.7. *Code Administration:* The Administrator of this Code is DIGI who has established an Administration Sub-committee comprising representatives from Signatories and independent members who will meet at six monthly intervals to review the actions of Signatories and monitor how they are meeting their commitments under the Code. In addition, DIGI has appointed an independent expert to fact check and attest Signatories’ annual transparency reports under the Code, in order to incentivise best practice and compliance.

- 7.8. *Code Review:* The Code will be reviewed at two yearly intervals. The reviews will be based on the input of the Signatories, and on relevant government bodies (including the Australian Communications and Media Authority) and other interested stakeholders including academics and representatives from civil society active in this field. Amendments to the Code will be published on the DIGI website.

APPENDIX 1

Opt-in Nomination Form

This form will be used by Signatories to indicate their commitments in accordance with this Code. Signatories should include a brief explanation as to why they have not elected to make some or all of the optional commitments in accordance with section 7.1.

Objective	Outcome	Measure		Opt-in/ Opt-out
Objective 1: Provide Safeguards against Harms that may arise from Disinformation and Misinformation.	Outcome 1a: Signatories contribute to reducing the risk of Harms that may arise from the propagation of Disinformation and Misinformation on digital platforms by adopting a range of scalable measures.	5.8	Signatories will develop and implement measures which aim to reduce the propagation of and potential exposure of users of their services and products to Disinformation and Misinformation	Yes
		5.9	Measures implemented under 5.8, may include, by way of example rather than limitation: a) policies and processes that require human review of user behaviours or content that is available on digital platforms (including review processes that are conducted in partnership with fact-checking organisations); b) labelling false content or providing trust indicators of content to users; c) demoting the ranking of content that may expose users to Disinformation; d) removal of content which is propagated by Inauthentic Behaviours e) notifying users when they have been exposed to Disinformation; f) suspension or disabling of accounts of users which engage in Inauthentic Behaviours; g) the provision or use of technologies to identify and reduce Inauthentic Behaviours that can expose users to Disinformation such as	

			<p>algorithmic review of content and/or user accounts;</p> <p>h) the provision or use of technologies which assist digital platforms or their users to check authenticity or accuracy or to identify the provenance or source of digital content</p> <p>i) exposing meta data to users about the source of content;</p> <p>j) enforcing published editorial policies and content standards;</p> <p>k) prioritising credible and trusted news sources that are subject to a published editorial code:</p> <p>l) partnering and/or providing funding for fact checkers to review Digital Content; and</p> <p>m) providing users with tools that enable them to exclude their access to certain types of Digital Content.</p>	
	Outcome 1b: Users will be informed about the types of behaviours and types of content that will be prohibited and/or managed by Signatories under this Code.	5.10	Signatories will implement and publish policies and procedures and any appropriate guidelines or information relating to the prohibition and/or management of user behaviours that may propagate Disinformation and Misinformation via their services or products.	
	Outcome 1c: Users can report content and behaviours to Signatories that violates their policies under 5.10 through publicly available and accessible reporting tools.	5.11	Signatories will implement and publish policies, procedures and any appropriate guidelines or information regarding the reporting of the types of content and behaviours that may propagate Disinformation and Misinformation via their platforms.	
		5.12	In implementing the commitment in 5.11 Signatories recognise that the terms Disinformation and Misinformation may be	

			unfamiliar to users and thus policies and procedures aimed at achieving this outcome may specify how users may report a range of impermissible content and behaviours on Digital platforms.	
	Outcome 1d: Users will be able to access general information about Signatories' actions in response to reports made under 5.11.	5.13	Signatories will implement and publish policies, procedures and/or aggregated reports (including summaries of reports made under 5.11) regarding the detection and removal of content that violates platform policies, including but not necessarily limited to content on their platforms that qualifies as Misinformation and/or Disinformation.	
	Outcome 1e: Users will be able to access general information about Signatories' use of recommender systems and have options relating to content suggested by recommender systems.	5.14	Signatories that provide services (other than search engines) whose primary purpose is to disseminate information to the public and which use recommender systems, commit to : A. make information available to end-users about how they work to prioritise information that end-users may access on these services; and B. provide end-users with options that relate to content suggested by recommender systems that are appropriate to the service.	
Objective 2: Disrupt advertising and monetisation incentives for Disinformation	Outcome 2: Advertising and/or monetisation incentives for Disinformation are reduced.	5.15	Signatories that offer digital advertising services will use commercially reasonable efforts to deter advertisers from repeatedly placing digital advertisements that propagate Disinformation or Misinformation.	
		5.16	Signatories will implement policies and processes that aim to disrupt advertising and/or	

			monetisation incentives for Disinformation.	
		5.17	<p>Policies and processes required under 5.16 may for example:</p> <ul style="list-style-type: none"> a) Promote and/or include the use of brand safety and verification tools; b) Enable engagement with third party verification companies; c) Assist and/or allow advertisers to assess media buying strategies and online reputational risks; d) Provide advertisers with necessary access to client-specific accounts to help enable them to monitor the placement of advertisements and make choices regarding where advertisements are placed; and /or e) restrict the availability of advertising services and paid placements on accounts and websites that propagate Disinformation. 	
		5.18	<p>Signatories recognise that all parties involved in the buying and selling of online advertising and the provision of advertising-related services need to work together to improve transparency across the online advertising ecosystem and thereby to effectively scrutinise, control and limit the placement of advertising on accounts and websites that propagate Disinformation.</p>	
Objective 3: Work to ensure the security and integrity of services and products delivered by Digital platforms.	Outcome 3: The risk that Inauthentic User Behaviours undermine the integrity and security of services and products is reduced.	5.19	<p>Signatories commit to take measures that prohibit or manage the types of user behaviours that are designed to undermine the security and integrity of their services and products, for example, the use of fake accounts or automated bots that are designed to propagate Disinformation.</p>	
		5.20	<p>To allow for the expectations of some users and Digital</p>	

			platforms about the protection of privacy, measures developed and implemented in accordance with this commitment should not preclude the creation of pseudonymous and anonymous accounts.	
Objective 4: Empower consumers to make better informed choices of digital content.	Outcome 4: Users are enabled to make more informed choices about the source of news and factual content accessed via digital platforms and are better equipped to identify Misinformation.	5.21	Signatories will implement measures to enable users to make informed choices about news and factual information and to access alternative sources of information.	
		5.22	Measures developed and implemented in accordance with the commitment in 5.21 may for example include: a) the use of technological means to prioritise or rank news content to enable users can easily find diverse perspectives on matters of public interest; b) aggregation or promotion of news content subject to an independent editorial code and complaints scheme; c) the provision or use of technologies which signal the credibility of news sources or which assist Digital platforms or their users to check the authenticity or accuracy of online news content or to identify its provenance or source d) the promotion of digital literacy initiatives informed by evidence or expert analysis; and or e) the provision of financial support for sustainable partnerships with fact-checking organisations.	
Objective 5: Improve public awareness of the source of Political Advertising	Outcome 5: Users are better informed about the source of Political Advertising.	5.23	Signatories will develop and implement policies that provide users with greater transparency about the source of Political Advertising carried on digital platforms.	

carried on digital platforms.		5.24	Measures developed and implemented in accordance with the commitment in 5.23 may include requirements that advertisers to identify and/or verify the source of Political Advertising carried on digital platforms; the provision of tools which enable the public to understand whether a political ad has been targeted to them; policies which require that Political Advertisements which appear in a medium containing news or editorial content are presented in such a way as to be readily recognisable as a paid-for communication.	
		5.25	Signatories may also, as a matter of policy, choose not to target advertisements based on the inferred political affiliations of a user.	
Objective 6: Strengthen public understanding of Disinformation and Misinformation through support of strategic research.	Outcome 6: Signatories support the efforts of independent researchers to improve public understanding of Disinformation and Misinformation.	5.26	Signatories commit to support and encourage good faith independent efforts to research Disinformation and Misinformation both online and offline. Good faith research includes research that is conducted in accordance with the ethics policies of an accredited Australian University provided such policies require that data collected by the researcher is used solely for research purposes and is stored securely on a university IT system or any research which is conducted in accordance with the prior written agreement of the digital platform.	
		5.27	Measures taken to implement 5.26 may include, for example, cooperation with relevant initiatives taken by independent fact checking bodies. Other measures may include funding for research and/or sharing datasets, undertaking joint research, or otherwise partnering with academics and civil society organisations.	

		5.28	Signatories commit not to prohibit or discourage good faith research into Disinformation and Misinformation on their platforms.	
		5.27	Relevant Signatories commit to convene an annual event to foster discussions regarding Disinformation within academia and Civil Society.	
Objective 7: Signatories publicise the measures they take to combat Disinformation and Misinformation.	Outcome 7: The public can access information about the measures Signatories have taken to combat Disinformation and Misinformation.	5.30	Signatories will make and publish a transparency report information as set out in section 7 (see 7.3 for the form of reports for larger platforms with a user base in excess of one million monthly active Australian end- users and section 7.4 for the form of reports for smaller platforms with a user base less than one million monthly active Australian end- user)	Yes
		5.31	In addition, Signatories will publish additional information detailing their progress in relation to Objective 1 and any additional commitments they have made under this Code.	
		5.32	Signatories may fulfill their commitment in section 5.31 by providing additional reports and/or public updates on areas such as content removals, open data initiatives, research reports, media announcements, user data requests and business transparency reports. Examples of such information could include, by way of example rather than limitation, blog posts, white papers, in-product notifications, transparency reports, help centres, or other websites.	

APPENDIX 2

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 mis/disinformation reporting under 1c?]

[Include such data if available]

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

Outcome 1e: Users can access general information about Signatories' use of recommender systems and have options relating to content suggested by recommender systems.

Objective 2: Disrupt advertising and monetisation incentives for disinformation.

[Explain KPIs as above]

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

[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 mis/disinformation 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]

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 mis/disinformation 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

Business and Content Context

Approach to Disinformation and Misinformation