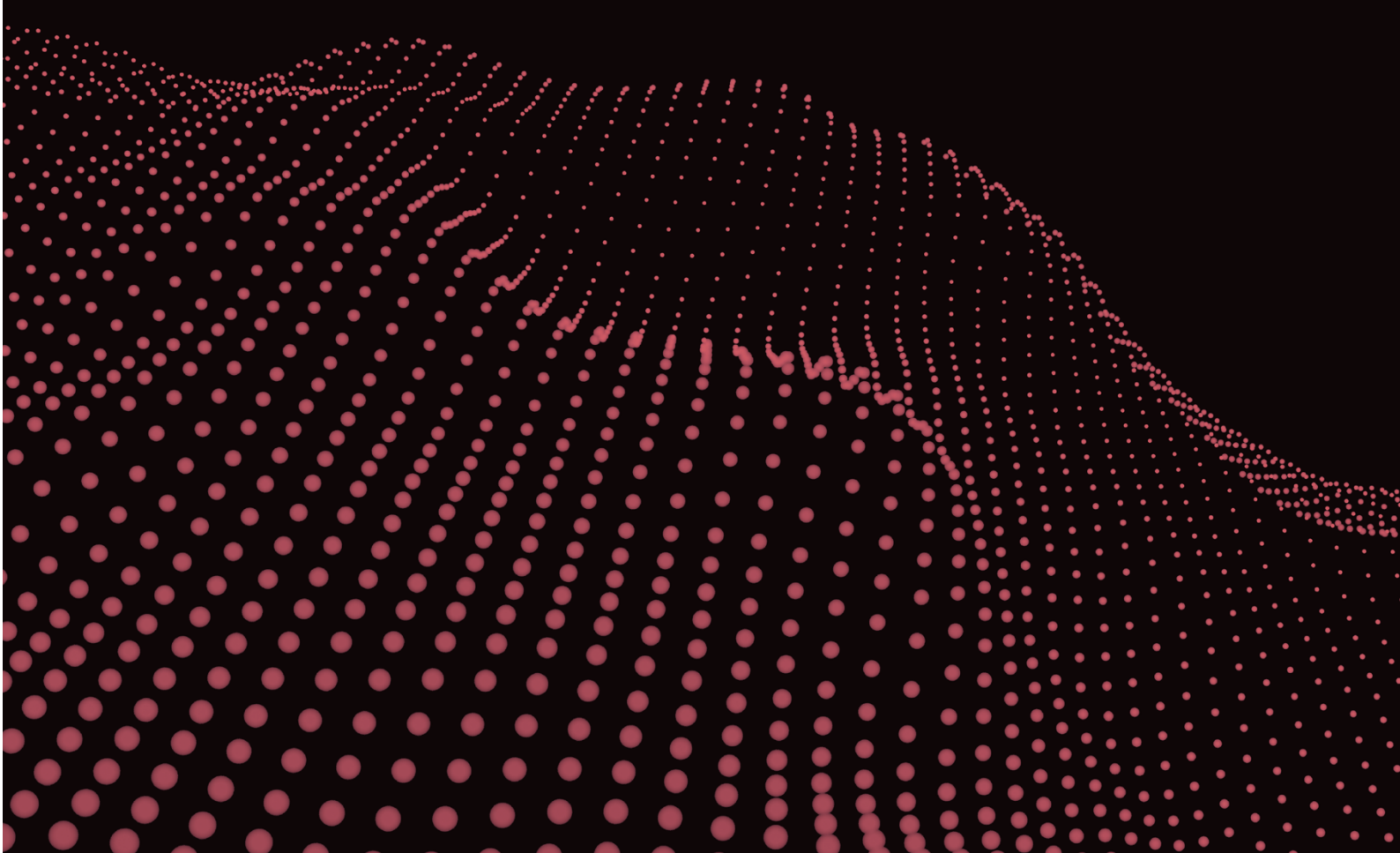




Australian Code of Practice on Disinformation and Misinformation

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

PUBLISHED FEBRUARY 22, 2021



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*.
- 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. *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.

- 1.5. *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.
- 1.6. *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.
- 1.7. *Signatories may take a more expansive approach:* Signatories may in their discretion implement policies and processes that contain a more expansive set of obligations than is provided for under this Code.
- 1.8. *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.
- 1.9. *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.
- 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 an imminent 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.

- 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. *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 or referendum; or
 - C. that advocate for the outcome of a political campaign concerning a social issue of public concern in Australia; or
 - D. are regulated as political advertising under Australian law.
- 3.8. *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 based 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.

4. Scope, application and commencement of this Code

4.1. *Scope:* Recognising that the types of user behaviour and content that is subject to the Code will vary greatly in incidence and impact amongst the diverse range of services and products offered by different digital platforms, it is expected that the commitments under this Code will apply to the services and products that deliver to end users in Australia:

- A. user-generated (including sponsored and shared) content: and/or
- B. content that is returned and ranked by Search Engines in response to user queries.

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 excluded services and products is not intended to be exhaustive as new services and products are likely to emerge, some of which will not be relevant to the Code.

4.4. *Excluded Content:* The following content is excluded from the operation of the Code unless Signatories determine that specific instances of its propagation clearly fall within the scope of Disinformation set out in section 3.2:

- 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 5.21 to 5.23, Political Advertising or content authorised by a political party registered under Australian law; and
- D. news content that is the subject of a published editorial code which sets out content standards and or/complaints mechanisms.

- 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 operation of the provisions of the Code 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 *Enhancing Online Safety Act 2015 (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 commences on 22 February 2021.

5. Objectives and Measures

- 5.1. *General:* This section incorporates a range of measures aimed at achieving seven key objectives and ten outcomes which are informed by the purpose and guiding principles of the Code 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 an annual report 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 meta data 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 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 any 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 violates 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.

Objective 2: Disrupt advertising and monetisation incentives for Disinformation.

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

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

5.15. Policies and processes implemented under 5.14 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.

5.16. 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.

- 5.17. 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.
- 5.18. 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.

- 5.19. Signatories will implement measures to enable users to make informed choices about Digital Content and to access alternative sources of information.
- 5.20. Measures developed and implemented in accordance with the commitment in 5.19 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; 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.

- 5.21. While Political Advertising is not Misinformation or Disinformation 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.
- 5.22. Measures developed and implemented in accordance with the commitment in 5.21 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.
- 5.23. 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.

- 5.24. 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.
- 5.25. Measures taken to implement 5.24 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.
- 5.26. Signatories commit not to prohibit or discourage good faith research, as described in 5.24, into Disinformation or Misinformation on their platform.
- 5.27. Relevant Signatories commit to convene an annual event to foster discussions regarding Disinformation and Misinformation 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.

- 5.28. All Signatories will make and publish the annual report information in section 7.
- 5.29. 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.

- 5.30. Signatories may fulfill their commitment in section 5.29 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.
- 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 Report:* In addition to the Opt-in Nomination Form under section 7.1, 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. The first report will be in the form of the template report provided in Appendix 2 and submitted within three months of the commencement of the Code. Signatories commit to develop and implement, within six months of the commencement of this Code, an agreed format for future annual reports and a guideline that will inform the data and other information to be included in subsequent reports.

- 7.4. *Complaints:* Signatories agree to establish, within six months of the commencement of this Code, a facility for addressing non-compliance by Signatories with their general commitments of the Code (including to the seven objectives and ten outcomes). The facility will hear appeals of complaints of Code breaches that have not been acted upon by Signatories (but not individual complaints of Signatories' decisions regarding content on their platforms, including whether specific items of content should be retained or removed). Signatories also agree to develop and document a process describing circumstances in which a non-compliant Signatory may be removed. 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.
- 7.5. *Code Administration:* The Administrator of this Code is DIGI who will establish a 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.
- 7.6. *Code Review:* The Code will be reviewed after it has been in operation for twelve months, and thereafter 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.

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 any 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: <ul style="list-style-type: none"> 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 	

			<p>as 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>	
	<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.</p>	5.10	<p>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.</p>	
	<p>Outcome 1c: Users can report content and behaviours to Signatories that violates their policies under 5.10 through publicly available and accessible reporting tools.</p>	5.11	<p>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.</p>	
		5.12	<p>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</p>	

			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.	
Objective 2: Disrupt advertising and monetisation incentives for Disinformation	Outcome 2: Advertising and/or monetisation incentives for Disinformation are reduced.	5.14	Signatories will implement policies and processes that aim to disrupt advertising and/or monetisation incentives for Disinformation.	
		5.15	Policies and processes required under 5.14 may for example: 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) E) restrict the availability of advertising services and paid placements on accounts and websites that propagate Disinformation.	
		5.16	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 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.	5.17	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.	
		5.18	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.	5.18	Signatories will implement measures to enable users to make informed choices about news and factual information and to access alternative sources of information.	
		5.19	Measures developed and implemented in accordance with the commitment in 5.18 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	

			<p>identify its provenance or source</p> <p>d) the promotion of digital literacy; and or</p> <p>e) E) the provision of financial support for sustainable partnerships with fact-checking organisations.</p>	
<p>Objective 5: Improve public awareness of the source of Political Advertising carried on digital platforms.</p>	<p>Outcome 5: Users are better informed about the source of Political Advertising.</p>	5.21	<p>Signatories will develop and implement policies that provide users with greater transparency about the source of Political Advertising carried on digital platforms.</p>	
		5.22	<p>Measures developed and implemented in accordance with the commitment in 5.21 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.</p>	
		5.23	<p>Signatories may also, as a matter of policy, choose not to target advertisements based on the inferred political affiliations of a user.</p>	
<p>Objective 6: Strengthen public understanding of Disinformation and Misinformation through support of strategic research.</p>	<p>Outcome 6: Signatories support the efforts of independent researchers to improve public understanding of Disinformation and Misinformation.</p>	5.24	<p>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</p>	

			the prior written agreement of the digital platform.	
		5.25	Measures taken to implement 5.24 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.	
		5.26	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.	5.28	All Signatories will make and publish the annual report information in section 7	Yes
		5.29	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.	
		5.30	Signatories may fulfill their commitment in section 5.29 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

[Name of Company]: Australian Code of Practice on Disinformation and Misinformation Initial Report

Executive summary

This Section should contain an overview of the findings of each Signatory's report with respect to the efforts they have made to implement measures to address each commitment they have made under the Code. It should also provide some observations about future measures, for example by detailing areas where they may be testing new approaches.

Background

Business and Content Context

Signatories should provide some broader context regarding the way the Code applies to their business such as the types of service they provide to users in Australia and the types of content they provide which is the subject of commitments under the code e.g., advertising content (including sponsored content), user generated content and search engine results. They should also set out any factors in section 6.1 of the code (proportionality) that they consider are relevant in assessing the measures taken under the Code.

Approach to Disinformation and Misinformation

In addition, Signatories should give some context regarding their approach to Disinformation and Misinformation including the extent to which their policies and procedures exceed the minimum standards required by the Code. They should also provide additional data about the way Disinformation and Misinformation is manifested on the products and services provided via platforms (what kinds of Harms and behaviors they are concerned about).

Approach to monitoring performance

Signatories should explain how they assess their performance relation to the commitments that have made under the Code e.g., efforts they are making to cooperate

with third parties such as fact-checking organisations or technology partners to combat Disinformation and Misinformation.

Data about trends

If available, it would be useful to provide some qualitative or quantitative data such as case studies and any comparative data sets that explain the nature and extent of the issue for users in Australia. Signatories should also include any insights they have about trends in Disinformation e.g., the types of actors involved and the types of inauthentic behaviors they have identified in Australia and globally which amplify Disinformation content.

Future Initiatives

Signatories may wish to provide some information about future measures they may take in furtherance of their commitments under the Code.

Objective 1: Safeguards against 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.

Signatories will describe specific actions they have taken to meet this commitment, including any relevant quantitative or qualitative data (including case study examples) about the outcomes of these measures in the Australian context.

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.

Signatories will include links to published policies and procedures, guidelines and 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 or behaviours to Signatories that violates their policies under section 5.10 through publicly available and accessible reporting tools.

Signatories will include links to published policies, procedures guidelines that will enable users to report the types of behaviours and content that may propagate Disinformation and Misinformation via their platforms.

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

Signatories will give details about how and when they have published general information about their responses to reports by users about content that violates their policies.

Objective 2: Disrupt advertising and monetisation incentives for disinformation.

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

Signatories will describe the policies, processes and products that they have developed and/or implemented in order to disrupt advertising and/or monetisation incentives for behaviours that may propagate Disinformation. This Section should also contain any relevant qualitative or quantitative data (such as case study examples) about the extent those reduce advertising and/or monetisation incentives for Disinformation in relation to content that is provided to users in Australia.

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

Signatories will describe the policies and processes that they have implemented that prohibit or manage the types of user behaviours that may undermine the integrity and

security of their services and products. This Section should also contain any relevant qualitative or quantitative data (including case study examples) about the extent those measures have reduced the risk that inauthentic user behaviours undermine the integrity and security of their services and products delivered to Australian users.

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.

Signatories detail measures implemented to enable users to make more informed choices about the source of news and factual content accessed via digital platforms and to access alternative sources. This Section should also contain any relevant qualitative or quantitative data about Australian users' response to measures such as the extent they have used empowerment tools (including case study examples) in relation to different categories of content (e.g., advertising, news, academic research, search engine results) provided on digital platforms.

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.

Signatories detail policies that provide users with information about the source of Political Advertising carried on digital platforms.

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.

Signatories detail measures to support for the efforts of independent researchers to improve the Australian public understanding of Disinformation and Misinformation both online and offline. This Section could for example set out how Signatories are cooperating with the research community including, for example, through funding of research through the provision of tools that facilitate the running of queries by researchers and fact-checkers and enabling independent monitoring and analysis of disinformation trends and assessment of the measures taken by platforms under this code. Signatories should also discuss issues (such as data protection concerns) that may have impacted on cooperation with the research community. This Section should also contain any relevant qualitative or quantitative data (including case study examples) about the extent those measures have assisted research into the experience of Disinformation in Australia.

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

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

Signatories list and/or provide links to reports made available to government and the public on relevant efforts under this Code.

Concluding remarks

This Section should contain any further observations Signatories wish to make about their response to the Code.