AUSTRALIAN CODE OF PRACTICE ON DISINFORMATION

1 Preamble

1.1 *Background:* This Code of Practice 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 is an aspect of a wider, multifaceted social problem. Manipulative actors have long engaged in a range of offline and online behaviours which propagate information that have threatened 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 amongst users of their Services and Products. This Code is principally designed to express the commitments made by Digital Platforms to address the propagation of Disinformation.

1.4 *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 the propagation of Disinformation 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 for their Digital Platform, Services and Products, this Code provides Signatories the ability to opt into a range of measures.

1.5 *Proportionality:* The 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. 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 kinds of contextual factors that Signatories may take into account in this regard.

1.6 *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 including public authorities, academia, civil society, and news organisations. Tackling Disinformation 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, what does not, and how together we can respond to the evolving challenges of Disinformation.

2 Guiding Principles

2.1 *The purposes of the Code:* Digital Platforms recognise the need to address the tactics and behaviours that are used to propagate Disinformation via their Services and Products. In developing the Code, the Signatories need to be clear about their perspective on the nature of the challenges posed by Disinformation in the context of the services and products they provide to Australian users. Therefore, the Signatories have set out in sections 2.2 to 2.8 a set of principles that have guided their design of the Code and its application to their Services and Products.

2.2 *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 perceived falsity if the content would not otherwise be unlawful. To that end, this Code is designed to enable Signatories to carefully balance their commitments to protect their users from potentially harmful content with the rights of users to engage in free speech and expression and the importance of enabling diverse perspectives and voices to be heard in political debate.

2.3 *Protection of user privacy:* Digital Platforms value their users' privacy. Any actions taken by Digital Platforms to address the propagation of Disinformation 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 law and regulations. 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.

2.4 *Scrutiny of advertising placement.* Digital Platforms recognise the importance of scrutinising advertisement placements on their Services and Products to reduce revenues that may reach the propagators of Disinformation.

2.5 *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 that may cause Harm.

2.6 *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

2.7 *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. The Code provides various options for Digital Platforms to participate in independent research initiatives.

2.8 *Without prejudice commitments* This Code is without prejudice to other initiatives aimed at tackling Disinformation by Digital Platforms.

3 Glossary

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

3.1 *Digital Platforms* under this Code include those Signatories with a presence in Australia who have signed this Code and have made commitments in respect of all or part of its provisions in relation to all or some of their Services and Products.

3.2 The aspect of *Disinformation* that this Code focuses on is cumulatively:

A) Inauthentic Behaviour by users of a Digital Platform;
B) that propagates digital content via that platform;
C) for the purpose of economic gain or to mislead or deceive the public;
D) that may cause Harm; and
E) is not otherwise unlawful.

Disinformation does not include misleading advertising, reporting errors, satire and parody, or clearly identified partisan news and commentary.

3.3 *Harm* means an imminent and serious threat to:

A) democratic political and policymaking processes; or
B) public goods such as the protection of citizens' health, the environment or security.

3.4 *Inauthentic Behaviour* includes spam and other forms of deceptive behaviours (including via automated systems) which encourages users of Digital Platforms to propagate content which may cause Harm.

3.5 *Enterprise Software* is software which is designed for use of a specific organisation.

3.6 *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 automatically collect information from and about internet sites. Second, they use sophisticated algorithms to return results in a set of links to websites, ranked based generally on relevance to the user’s specific search 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.

3.7 *Services and Products* means those services and or products that are made available by Digital Platforms which are subject to the Code as further explained in 4.1and 4.2.

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 content, and / or
B) content that is selected 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 Software;
D) content including e-books, videos, films, television, radio broadcasts or podcasts that is provided for the purpose of entertainment or education; and
E) content that is posted online that is authorised by an Australian State or Federal Government.

4.3 The list of excluded Services and Products is not intended to be exhaustive as new Products and Services are likely to emerge, some of which will not be relevant to the Code

4.4 *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 and election advertising) 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.

4.5 *Application*: The commitments made by each Signatory apply to it, in respect of the commitments it adopts, from the date that it opts into those commitments.

4.6 *Commencement*: This Code commences on [insert date].

5 Measures

5.1 *General:* This section incorporates a range of measures aimed at achieving six key objectives and eight outcomes which are informed by the purpose and guiding principles of the Code set out in section 2 above.

5.2 *Signatories*. 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 6 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 and the provision of reports may use terminology other than "Disinformation" and may, for example, refer to "Misinformation" 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 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 relating to the potential exposure of users to Disinformation, 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 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 or which increase its virality.

Objective 1: Safeguards against Disinformation

Outcome 1a: Signatories implement policies, processes and technologies which aim to reduce the risk that users of their Services and Products will be exposed to Disinformation.

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.

5.9 Measures taken by Signatories pursuant to section 5.8 may, for example, include policies and processes that require human review of user behaviours or content that is viral on Digital Platforms (including review processes that are conducted in partnership with fact-checking organisations) and/or the provision or use of technologies to identify and address behaviours that can expose users to Disinformation or which assist Digital Platforms or their users to check authenticity or accuracy or to identify the source of online content.

Outcome 1b: Signatories inform users about behaviours that will be prohibited and/or managed in order to limit users' exposure to Disinformation via their platforms.

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 via their Services or Products.

Outcome 1c: Signatories implement policies and procedures that enable users to report behaviours that may propagate Disinformation and to make these policies publicly available and accessible to users of their Services and Products.

5.11 Signatories will implement and publish policies, procedures and any appropriate guidelines or information regarding the reporting of the types of behaviours that may propagate Disinformation via their platforms.

5.12 In implementing the commitment in 5.11 Signatories recognise that the term Disinformation 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 behaviours on Digital Platforms.

Objective 2: Disrupt advertising and monetisation incentives for Disinformation

Outcome:**2**: Signatories implement policies and processes that aim to disrupt advertising and/or monetisation incentives for Disinformation.

5.13 Signatories will implement policies and processes that aim to disrupt advertising and/or monetisation incentives for behaviours that may propagate Disinformation.

5.14 Policies and processes required under 5.13 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) restrict the availability of advertising services and paid placements on accounts and websites that propagate Disinformation.

5.15 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 public benefit of Services and Products delivered by Digital Platforms

Outcome 3: Signatories implement measures aimed at protecting the public benefit of their Services and Products.

5.16 Signatories commit to take measures that prohibit or manage the types of user behaviours that are designed to undermine the public benefit of their Products and Services, for example, the use of fake accounts or automated bots that are designed to propagate Disinformation.

5.17 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: Signatories implement measures that enable users to make more informed choices about the source of news and factual content accessed via Digital Platforms that concerns matters which may cause Harm.

5.18 Signatories will implement measures to enable users to make informed choices about news and factual information that concerns matters which may cause Harm, and to access alternative sources of information on these matters. Measures developed and implemented in accordance with the commitment may include: the use of technological means to prioritise or rank content in search or news feeds or to provide ways that users can easily find diverse perspectives on matters of public interest; aggregation or creation of content subject to an editorial code; or 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 content or to identify its source.

Objective 5: Strengthen public understanding of Disinformation through support of strategic research

Outcome 5: Signatories provide support for the efforts of independent researchers to improve public understanding of Disinformation. Signatories commit to support and encourage good faith independent efforts to research Disinformation both online and offline.

5.19 Measures taken to achieve Objective 5 include, for example, cooperation with relevant initiatives taken by independent fact checking bodies. Other measures may include funding for research and/or sharing privacy protected datasets, undertaking joint research, or otherwise partnering with academics and civil society organisations.

5.20 Signatories commit not to prohibit or discourage good faith research into Disinformation on their platforms.

5.21 Relevant Signatories commit to convene an annual event to foster discussions regarding Disinformation within academia and Civil Society.

Objective 6: Signatories will publicise the measures they take to combat Disinformation

Outcome 6: Signatories provide reports to government and the public on relevant efforts under this Code**.**

5.22 Signatories will make and provide annual reports or updates to government and/or the public detailing their progress in relation to their commitments under this Code.

5.23 Signatories may fulfill their commitment in 5.22 by supplementing and drawing on information Signatories voluntarily provide as part of their external communications efforts and which are contained within a variety of 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.

6 Guidance on platform-specific measures

6.1 **Proportionality of measures:** The measures taken by Signatories pursuant to this Code will be proportionate and relevant to their specific context. Signatories may take into consideration a variety of factors including:

A) the severity and incidence of the behaviour concerned;
B) the actors which are engaged in the behaviour;
C) the nature of the behaviour, for example, whether it is intentional and/or maliciously motivated;
D) the breadth and speed by which content is disseminated;
E) the channel via which the content is distributed;
F) the nature of the content that is being propagated and the extent to which it is verifiably false;

G) the nature and extent of the Harm that is reasonably likely to result from the behaviour;
H) the type of Product or Service on which the behaviour occurs;
I) the size and nature of the Digital Platform's business and the resources available to it;
J) the need to protect freedom of expression; and
K) the need to protect user privacy.

6.2 **Examples of measures:** Examples of specific measures that might be adopted under this Code are set out below, noting that this list is a guide only and is not exhaustive nor relevant to all Digital Platforms.

A) complaints handling systems and processes;
B) flagging or demoting or ranking of content;
C) removal of content
D) user communication;
E) security systems;
F) account suspension or disabling;
G) technological and algorithmic review of content and/or accounts;
H) notifying users who have engaged with relevant content;
I) fact checking;
J) exposing meta data to users about the source of content;
K) partnerships with third-party organisations;
L) preventing monetisation of disinformation;
M) prioritising credible and trusted news sources;
N) editorial and curation processes.

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. Code Signatories will nominate the provisions of this Code 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.

7.1 *Annual Report*. Each signatory will provide an annual report to DIGI setting out its progress in relation to the outcomes contained in the Code. The reports will be published on [the DIGI] website. The first annual update will be submitted within three months of the anniversary of the commencement of the Code, with subsequent reports due at 12 monthly intervals. Signatories commit to develop and implement, within six months of the commencement of this Code, an agreed format for annual reports and a guideline that will inform the data and other information to be included in the reports.

7.2 *Complaints.* Signatories agree to establish, within six months of the commencement of this Code, a facility for addressing complaints about compliance with the Code. 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.

7.3 *Code Administration.* The Administrator of this Code is [DIGI] who will establish a sub-committee who will meet at six monthly intervals to review the actions of Signatories and monitor how they are meeting their commitments under the Code.

7.4 *Code Review*. The Code will be reviewed after it has been in operation for two years. The review will be based on the input of the Signatories, and on relevant government bodies 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.

Objective	Outcome	Measure		Opt-in
Objective 1: Safeguards against Disinformation	**Outcome 1a**: Signatories implement policies, processes and technologies which aim to reduce the risk that users of their Services and Products will be exposed to Disinformation.	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.	
		5.9	Measures, taken pursuant by Signatories pursuant to section 5.8 may, for example, include policies and processes that require human review of user behaviours or online content that is viral on Digital Platforms (including review processes that are conducted in partnership with fact-checking organisations) and/or the provision or use of technologies to identify and address behaviours that can expose users to Disinformation or which assist Digital Platforms or their users to check the authenticity or, accuracy or to identify the source of online content.	
	Outcome 1b: Signatories inform users about behaviours that will be prohibited and/or managed in order to limit users' exposure to Disinformation via their platforms.	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 via their Services or Products.	
	Outcome 1c: Signatories implement policies and procedures that enable users to report behaviours that may propagate Disinformation and to make these policies publicly available and accessible to users of their Services and Products.	5.11	Signatories will implement and publish policies, procedures and any appropriate guidelines or information regarding the reporting of the types of behaviours that may propagate Disinformation via their platforms.	
		5.12	In implementing the commitment in 5.11 Signatories recognise that the term Disinformation 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 behaviours on Digital Platforms.	
Objective 2: Disrupt advertising and monetisation incentives for Disinformation	**Outcome 2:** Signatories implement policies and processes that aim to disrupt advertising and/or monetisation incentives for Disinformation.	5.13	Signatories will implement policies and processes that aim to disrupt advertising and/or monetisation incentives for behaviours that may propagate Disinformation.	
		5.14	Policies and processes required under 5.13 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) restrict the availability of advertising services and paid placements on accounts and websites that propagate Disinformation.	
		5.15	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 public benefit of Services and Products delivered by Digital Platforms.	**Outcome 3:** Signatories implement measures aimed at protecting the public benefit of their Services and Products.	5.16	Signatories commit to take measures that prohibit or manage the types of user behaviours that are designed to undermine the public benefit of their Products and Services, for example, the use of fake accounts or automated bots that are designed to propagate Disinformation.	
		5.17	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:** Signatories implement measures that enable users to make more informed choices about the source of news and factual content accessed via Digital Platforms that concerns matters which may cause Harm.	5.18	Signatories will implement measures to enable users to make informed choices about news and factual information that concerns matters which may cause Harm, and to access alternative sources of information on these matters. Measures developed and implemented in accordance with the commitment may include: the use of technological means to prioritise or rank content in search or news feeds or to provide ways that users can easily find diverse perspectives on matters of public interest; aggregation or creation of content subject to an editorial code; or 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 content or to identify its source.	

Objective 5: Strengthen public understanding of Disinformation through support of strategic research	**Outcome 5:** Signatories provide support for the efforts of independent researchers to improve public understanding of Disinformation. Signatories commit to support and encourage good faith independent efforts to research Disinformation both online and offline.	5.19	Measures taken to achieve Objective 5 include, for example, cooperation with relevant initiatives taken by independent bodies such as the network of fact checkers facilitated by the European Commission. Other measures may include funding for research and/or sharing privacy protected datasets, undertaking joint research, or otherwise partnering with academics and civil society organisations.	
		5.20	Signatories commit not to prohibit or discourage good faith research into Disinformation on their platforms.	
		5.21	Relevant Signatories commit to convene an annual event to foster discussions regarding Disinformation	

			within academia and Civil Society.	
Objective 6: Signatories will publicise the measures they take to combat Disinformation.	**Outcome 6:** Signatories provide reports to government and the public on relevant efforts under this Code.	5.22	Signatories will make and provide annual reports or updates to government and/or the public detailing their progress in relation to their commitments under this Code.	
		5.23	Signatories may fulfill their commitment in 5.22 by supplementing and drawing on information contained within a variety of 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.	