

# Ensuring a strong and meaningful Code on Disinformation

Submission to the Australian Code of  
Practice on Disinformation

---

**Jordan Guiao**

Associate Fellow, Centre for Responsible Technology

**November 2020**

## About The Australia Institute

The Australia Institute is an independent public policy think tank based in Canberra. It is funded by donations from philanthropic trusts and individuals and commissioned research. We barrack for ideas, not political parties or candidates. Since its launch in 1994, the Institute has carried out highly influential research on a broad range of economic, social and environmental issues.

## About the Centre for Responsible Technology

The Australia Institute established the Centre for Responsible Technology to give people greater influence over the way technology is rapidly changing our world. The Centre will collaborate with academics, activists, civil society and businesses to shape policy and practice around network technology by raising public awareness about the broader impacts and implications of data-driven change and advocating policies that promote the common good.

## Our philosophy

As we begin the 21st century, new dilemmas confront our society and our planet. Unprecedented levels of consumption co-exist with extreme poverty. Through new technology we are more connected than we have ever been, yet civic engagement is declining. Environmental neglect continues despite heightened ecological awareness. A better balance is urgently needed.

The Australia Institute's directors, staff and supporters represent a broad range of views and priorities. What unites us is a belief that through a combination of research and creativity we can promote new solutions and ways of thinking.

## Our purpose - 'Research that matters'

The Institute publishes research that contributes to a more just, sustainable and peaceful society. Our goal is to gather, interpret and communicate evidence in order to both diagnose the problems we face and propose new solutions to tackle them.

The Institute is wholly independent and not affiliated with any other organisation. Donations to its Research Fund are tax deductible for the donor. Anyone wishing to donate can do so via the website at <https://www.tai.org.au> or by calling the Institute on 02 6130 0530. Our secure and user-friendly website allows donors to make either one-off or regular monthly donations and we encourage everyone who can to donate in this way as it assists our research in the most significant manner.

Level 1, Endeavour House, 1 Franklin St

Canberra, ACT 2601

Tel: (02) 61300530

Email: [mail@tai.org.au](mailto:mail@tai.org.au)

Website: [www.tai.org.au](http://www.tai.org.au)

ISSN: 1836-9014

# Introduction

In December 2019, the federal government asked digital platforms to develop a voluntary code of practice to “counter disinformation and help users to better identify the quality of online news and information”, and asked the Australian Communications and Media Authority (ACMA) to oversee the code’s development.

In June 2020, ACMA released a discussion paper on misinformation and news quality which outlines what ACMA thinks the code should cover.

The digital platforms have assigned responsibility for developing the code to the Digital Industry Group Inc (DIGI), which is the industry group for large technology companies in Australia. In October, DIGI released a discussion paper and a draft industry code of practice.

The Centre for Responsible Technology supports the development of an Australian Code of Practice for Disinformation and believes it is a critical step in combating harmful disinformation in Australia.

A functioning Australian Code of Practice for Disinformation requires the large technology companies, whose platforms host, distribute and amplify disinformation, to take responsibility and treat disinformation seriously.

The Centre’s submission considers DIGI’s discussion paper and draft Code in the broader context of the disinformation landscape, and presents several critiques of the DIGI paper and Code.

DIGI’s “business as usual” approach does not meet the expectations and scope outlined by ACMA’s discussion paper.

This submission presents six recommendations for the development of the Code.

# Response to the discussion paper

The DIGI discussion paper clarifies the body's overall approach and highlights elements of the disinformation landscape, like malicious actors and the role of news. The paper also describes initiatives digital platforms like Google and Facebook have taken to date.<sup>1</sup>

Some platforms, notably including Twitter, appear to take the issue seriously with their recent policy updates around the US election.<sup>2</sup>

However, the DIGI discussion paper attempts to shift responsibility from the digital platforms by blaming broader issues and emphasising individual user action. The Code is specifically aimed at the digital platforms, and should require details for *their* specific responses.

## Self-regulation and administration of the Code

DIGI is an Australian lobby group for the “Big Tech” digital industry, including Google, eBay, Twitter and Facebook.<sup>3</sup> The federal government asked digital platforms to collaborate on an industry-wide standard, and facilitating this through a representative body makes sense.

However, the discussion paper and draft Code indicate that the platforms are taking a “business as usual” approach to the Code development.

A key criterion ACMA will use to determine the efficacy of the Code development is the robustness of consultation conducted by DIGI.<sup>4</sup> DIGI should reveal which organisations and experts they have consulted. Civic society, academics, human rights groups and other businesses could provide a different perspective to technology companies.

Self-regulation and self-reporting are not sufficient to ensure technology companies act on disinformation. Large digital platforms like Google and Facebook have benefitted significantly from a deregulated marketplace and this has allowed disinformation to spread. An exercise aimed at ensuring disinformation is addressed properly should not be facilitated through a tech-funded industry group.

---

<sup>1</sup> DIGI (2020) *Discussion Paper on an Australian Voluntary Code of Practice for Disinformation*, <https://digi.org.au/wp-content/uploads/2020/10/Discussion-Paper-Final.pdf>

<sup>2</sup> Gadde & Beykpour (2020) *Additional steps we're taking ahead of the 2020 US Election*, [https://blog.twitter.com/en\\_us/topics/company/2020/2020-election-changes.html](https://blog.twitter.com/en_us/topics/company/2020/2020-election-changes.html)

<sup>3</sup> DIGI (2020) *About DIGI*, <https://digi.org.au/about/>

<sup>4</sup> ACMA (2020) *Misinformation and news quality on digital platforms in Australia, a position paper to guide code development*, pg. 29, <https://www.acma.gov.au/sites/default/files/2020-06/Misinformation%20and%20news%20quality%20position%20paper.pdf>

In 2019, the French government’s inter-departmental “mission” on how to make social media platforms more accountable found that:

None of the information made public by the platforms concerning their self-regulatory actions can be corroborated by objective facts.<sup>5</sup>

Similarly, last year the former head of Facebook Australia Stephen Scheeler said:

We can't leave the tech giants to regulate themselves completely. The reason is I don't think they're capable of doing so.

Most companies aren't capable of self-regulation because there is an agency problem. Unless there are negative consequences, then you don't take things seriously. Negative consequences drive behaviour and focus the mind.<sup>6</sup>

To date, digital platform self-regulation regarding disinformation has been reactive, often occurring only after the damage has been done. Facebook’s Oversight Board – their single largest effort to address difficult content policy decisions – was announced in 2018 but only began accepting cases two weeks before the 2020 US election.<sup>7</sup>

For the Code to be a significant development in the fight against disinformation in Australia, it should be developed and facilitated by an independent body, not by the digital platforms or their lobby group.

## Hiding complicity in complexity

The DIGI discussion paper says that finding “salient and meaningful terms to define the issues remain complex and difficult”.<sup>8</sup> It also introduces the term “information disorder” to describe a “new” type of condition in the disinformation landscape. The paper also argues that disinformation solutions need to come from various groups, including media organisations and individual education and literacy.<sup>9</sup>

While the digital landscape can be complex, the purpose of *this* Code of Practice is to determine the disinformation response from digital platforms. This is not an enquiry on

---

<sup>5</sup> DINUM (2020) *Creating a French framework to make social media platforms more accountable: Acting in France with a European vision*, pg. 12, [https://www.numerique.gouv.fr/uploads/Regulation-of-social-networks\\_Mission-report\\_ENG.pdf](https://www.numerique.gouv.fr/uploads/Regulation-of-social-networks_Mission-report_ENG.pdf)

<sup>6</sup> Redrup & Tillery (2020) *Social media platforms can't self-regulate*, <https://www.afr.com/technology/social-media-platforms-can-t-self-regulate-20190327-p517y5>

<sup>7</sup> Wong (2020) *Facebook's long-awaited oversight board to launch before US election*, <https://www.theguardian.com/technology/2020/sep/24/facebook-oversight-board-launch-us-election>

<sup>8</sup> DIGI (2020) *Discussion Paper on An Australian Voluntary Code of Practice for Disinformation*, pg. 8, <https://digi.org.au/wp-content/uploads/2020/10/Discussion-Paper-Final.pdf>

<sup>9</sup> Ibid.

news organisations, or an audit of digital literacy. Technology companies generate revenue from disinformation<sup>10</sup> and are responsible for addressing it on their platforms.

It is a myth that digital platforms are too big or complex to fight disinformation.<sup>11</sup> The failure to address disinformation is not a result of scale or complexity, but a failure of imagination on the part of technology platforms and unwillingness to change their business models.

Platforms notoriously underinvest in (human) content moderators, with harrowing reports of moderators experiencing burnout and post-traumatic stress disorder (PTSD).<sup>12</sup> Facebook recently announced new Artificial Intelligence (AI) tools to combat disinformation, including an AI model architecture called Linformer, a deep fake detection tool, and new models to identify and classify hate speech.<sup>13</sup> But AI has limitations, with a New York University study finding that Facebook's AI made approximately 300,000 mistakes in content moderation each day.<sup>14</sup>

Until large digital platforms face consequences for failing to address disinformation, they will continue to underinvest in the necessary disinformation response.

Moreover, the solutions discussed by digital platforms often only address the downstream elements of disinformation, like content publishing and moderation, rather than root causes, which include the in-built product design of online platforms that provide velocity and reach to provocative content and generate money from its virality.

Tech exceptionalism, the idea that technology innovates so quickly that it is “too complex” for lay people and legislators, and should therefore not be bound by the same rules as other industries, is unacceptable.<sup>15</sup>

Complexity does not excuse complicity. Online companies need to take responsibility for the disinformation harms facilitated by their platforms.

---

<sup>10</sup> Paul (2020) *Youtube profits from videos promoting unproven Covid-19 treatments*, <https://www.theguardian.com/technology/2020/apr/03/youtube-coronavirus-treatments-profit-misinformation>

<sup>11</sup> Mims (2020) *'Too Complex to Break Up' is the new 'Too Big to Fail'*, <https://www.wsj.com/articles/too-complex-to-break-up-is-the-new-too-big-to-fail-11602302406>

<sup>12</sup> Dvoskin (2020) *Facebook content moderator details trauma that prompted fight for \$52 million PTSD settlement*, <https://www.washingtonpost.com/technology/2020/05/12/facebook-content-moderator-ptsd/>

<sup>13</sup> Wiggers (2020) *Facebook's improved AI isn't preventing harmful content from spreading*, <https://venturebeat.com/2020/11/19/facebooks-improved-ai-isnt-preventing-harmful-content-from-spreading/>

<sup>14</sup> Barrett (2020) *Who Moderates the Social Media Giants? A Call to End Outsourcing*, <https://bhr.stern.nyu.edu/tech-content-moderation-june-2020>

<sup>15</sup> Wagstaff (2020) *Tech exceptionalism is over, and now things need to change*, <https://sea.mashable.com/tech/8565/tech-exceptionalism-is-over-and-now-things-need-to-change>

## European guidelines are not the Australian guidelines

The European Code of Practice is a useful reference point in developing a similar code in Australia.

The discussion paper by DIGI notes that intended signatories for the Australian Code have in most cases already signed up to the European Code. This effort is welcome.<sup>16</sup> However, involvement in the European Code should not be used as an excuse to limit action in Australia.

While there will be overlaps between the Codes, the digital platforms must demonstrate that they are addressing the concerns of Australian governments and the Australian people.

The draft Australian Code fails to specifically addressing the terms of the Australian paper. Instead, it presents existing policies developed by digital platforms. The digital platforms must specifically address Australian requirements.

The Centre for Responsible Technology also notes that the United Kingdom has recommended a statutory duty of care, rather than a voluntary Code. The UK government rejected a Code-based approach because only a small number of companies engage, and participants will likely not do enough “to make systematic improvements in governance and risk management processes ... necessary”.<sup>17</sup> Clearly, the European Code is not considered adequate for all countries.

## The “freedom of speech” fallacy

Digital platforms sometimes argue that taking action against disinformation would breach a right to “freedom of speech”. This has been used by digital platforms as an excuse to resist taking acting on disinformation.

In Australia, digital platforms are not legally constrained from taking action on disinformation. Australia’s constitutional “implied freedom of political communication” is a limit on the legislative power of Parliament to make laws,<sup>18</sup> not a broad restriction on companies.

---

<sup>16</sup> DIGI (2020) *Discussion Paper on An Australian Voluntary Code of Practice for Disinformation*, <https://digi.org.au/wp-content/uploads/2020/10/Discussion-Paper-Final.pdf>

<sup>17</sup> UK Secretary of State for Digital, Culture, Media & Sport and the Secretary of State for the Home Department (2019) *Online Harms White Paper*, [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/793360/Online\\_Harms\\_White\\_Paper.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/793360/Online_Harms_White_Paper.pdf); (2020) *Consultation response*, <https://www.gov.uk/government/consultations/online-harms-white-paper/public-feedback/online-harms-white-paper-initial-consultation-response>

<sup>18</sup> Australian Human Rights Commission (2020) *Freedom of expression and the Internet*, <https://humanrights.gov.au/our-work/3-freedom-expression-and-internet>

Online platforms have no obligation to host disinformation and should not play into the hands of those who misuse “free speech” so that they can cause and incite harm against others.

The “freedom of speech” argument also glosses over the fact that digital platforms are structurally designed to allow content, including disinformation, to propagate widely and quickly.

Tackling this design in online platforms would be a much more meaningful way to combat disinformation so that the information landscape can be more manageable and disinformation is interrupted. One way to do this is to create more friction in online platform systems (as discussed below).

## An alternative disinformation “ABC” framework

The DIGI discussion paper suggests an “ABC” framework for disinformation:

- A: Manipulative actors
- B: Deceptive behaviour
- C: Harmful content [made by the actors]<sup>19</sup>

This matches DIGI’s strategy of emphasising the responsibility of others instead of focusing on the role digital platforms play.

The Centre for Responsible Technology proposes an alternative framework, which assigns responsibility to the platforms:

- A: Acceleration
- B: Black box
- C: Harmful content

### **A: Acceleration**

The speed and velocity built into the design of the online platforms contributes to the spread of disinformation. Content that elicits an emotional response, including a negative one, is more likely to receive mass engagement, and this is taken as a signal by technology algorithms to share this content faster and wider.

This has allowed fringe views, once restricted to a select group of sympathisers, to spread. Dangerous and inaccurate lies, conspiracy theories and hoaxes become mainstream as they are rewarded by algorithms for triggering engagement.

---

<sup>19</sup> DIGI (2020) *Discussion Paper on An Australian Voluntary Code of Practice for Disinformation*, <https://digi.org.au/wp-content/uploads/2020/10/Discussion-Paper-Final.pdf>



Introducing “friction” to deliberately interrupt velocity and virality can help curb the spread of disinformation. Friction is any step, content, feature or process of an online platform that creates an opportunity for users to pause, reflect and consider their behaviour.

The Forum on Information & Democracy’s policy framework on “infodemics” identifies numerous methods for creating software friction, including:<sup>20</sup>

- Providing more contextual information about the user who posted a content piece, including location, relationship to reader, affiliations, or expertise
- Providing information about how the content has been shared
- Sharing corrections to every user affected by a post which has been confirmed as disinformation
- Clarifying attributes that recommendation engines and algorithms use to target users’ specific content
- Cooling-off periods for targeted advertising
- Circuit breakers for viral content, like automatic halting once certain reach thresholds are met, and a required review of content
- A mandatory level of noise and randomness in content algorithms
- Limiting microtargeting
- Giving users more choices about how they are targeted

By introducing friction, disinformation’s acceleration can be disrupted at key points, and its virality reduced.

### **B: Black Box (Algorithmic opacity)**

Disinformation can spread because the engagement it elicits is rewarded by algorithms which do not distinguish between positive and negative interactions. Negative reactions are given the same weight and count towards the success of a content piece.

This is clearly a flaw in the design of algorithmic amplification from online platforms. Online platforms need to be more transparent in how their algorithms promote content, and algorithms that do not distinguish between useful and harmful content should be changed.

Three core functions need to be addressed:

- Ranking: How content is organised and ranked
- Targeting: How content is pushed to specific users and user groups
- Socialising: How user networks are used to influence and recommend specific actions and content<sup>21</sup>

---

<sup>20</sup> Forum on Information & Democracy (2020) *Working Group on Infodemics Policy Framework Report*, [https://informationdemocracy.org/wp-content/uploads/2020/11/ForumID\\_Report-on-infodemics\\_101120.pdf](https://informationdemocracy.org/wp-content/uploads/2020/11/ForumID_Report-on-infodemics_101120.pdf)

Most users do not want disinformation in their feeds; they did not seek it out. It is presented to them as a result of unknown attributes and variables which algorithms have decided for them – a significant information asymmetry between the user and the algorithm. The Forum on Information & Democracy finds:

A user risks becoming a vulnerable party in cases where they make detrimental choices that they otherwise would not have made were it not for the intervention of an algorithm. In such cases, users may experience a manifest disadvantage from ensuing actions that can be traced back to the dominating effects of a social platform's algorithm.<sup>22</sup>

To meaningfully combat disinformation, online platforms must make their algorithms and their functions more transparent so that problems can be identified and solutions found.

### **C: Harmful content**

Harmful content is at the core of disinformation. The challenge is ensuring that there are barriers to the spread of harmful content.

Instead of DIGI's emphasis on harmful content as created by harmful actors, the focus should be on the role digital platforms play in distributing and amplifying harmful content.

---

<sup>21</sup> Forum on Information & Democracy (2020) *Working Group on Infodemics Policy Framework Report*, pg. 21, [https://informationdemocracy.org/wp-content/uploads/2020/11/ForumID\\_Report-on-infodemics\\_101120.pdf](https://informationdemocracy.org/wp-content/uploads/2020/11/ForumID_Report-on-infodemics_101120.pdf)

<sup>22</sup> Forum on Information & Democracy (2020) *Working Group on Infodemics Policy Framework Report*, pg. 67, [https://informationdemocracy.org/wp-content/uploads/2020/11/ForumID\\_Report-on-infodemics\\_101120.pdf](https://informationdemocracy.org/wp-content/uploads/2020/11/ForumID_Report-on-infodemics_101120.pdf)

# Recommendations for the draft Code

ACMA has outlined clear terms of engagement for the development of the voluntary Australian Code of Practice for Disinformation.

The development of this Code was prompted by the recognition of the significance of the issue of disinformation and the key role that digital platforms play.

The current draft Code developed by DIGI is an inadequate response in the development of an appropriate Code of Practice.

The DIGI draft code reads like a “statement of intent” rather than a meaningful, tangible and quantifiable response. The draft Code’s Objectives are missing the necessary details.

Disinformation is a serious issue, and must be taken seriously by digital platforms.

At a minimum, the Centre for Responsible Technology recommends that the development of the Code of Practice include:

1. Specific examples and detail of the actions digital platforms are going to take against each Objective, not just statements of intent
2. Material that specifically addresses the Australian inquiry and the Australian terms of engagement.
3. Information from digital platforms on how they will introduce friction into the design of their products.
4. A commitment to transparency around algorithms from digital platforms, including details on how the mechanisms for reach and virality work.
5. The establishment of an independent body to facilitate ongoing developments and reports on progress, instead of an industry lobby group.
6. Contingencies, including stronger enforcement, should efforts to develop the Code be unsatisfactory or if the Code is not adhered to. Currently the Code is opt-in, and signatories can withdraw at any time.

# Conclusion

The Australian Voluntary Code of Practice for Disinformation is a critical step in tackling the harmful effects of disinformation in online platforms. This Code is part of an overall effort to correct the market imbalance caused by the dominance of certain digital platforms like Google and Facebook and resulting issues like the spread of disinformation.

ACMA has facilitated good faith developments of a voluntary industry Code and DIGI, the large technology companies' lobby group, has developed a draft paper and draft Code in response.

DIGI's response is not in the spirit of the original concerns and is an inadequate effort for combating urgent and harmful disinformation.

The "business as usual" approach and the "statement of intent" format of the draft Code do not satisfactorily meet expectations for Code development.

The Centre for Responsible Technology has critiqued several aspects of DIGI's response and made six recommendations which should be considered during the final Code development, including calling for more detail, introducing friction and algorithmic transparency, and the involvement of an independent body not funded by the large technology companies.