# Submission to DIGI

on the proposed

# Industry Disinformation Code

25 November 2020

DIGITAL RIGHTS WATCH

# Overview

We welcome the opportunity to submit comments to Digital Industry Group Inc (DIGI) on the draft industry disinformation code.

Disinformation is a large and complex topic and tackling it has become increasingly urgent task. Digital Rights Watch believes it is difficult to treat this problem as an issue of content governance or moderation alone, rather, as an issue endemic to certain advertising models, a lack of algorithmic transparency, and data collection/monetization practices.

Any proposed solution to the threat that the viral spread of misinformation poses to our democracies and the public marketplace of ideas on the Internet, must be underpinned by accountability and transparency in order to remain credible and instil lasting change.

## Digital Rights Watch

Digital Rights Watch is a charity organisation founded in 2016 whose mission is to ensure that people in Australia are equipped, empowered and enabled to uphold their digital rights. We stand for Privacy, Democracy, Fairness & Freedom in a digital age. We believe that digital rights are human rights which see their expression online.

We educate, campaign, and advocate for a digital environment where individuals have the power to maintain their human rights.[1]

# General remarks

### *Definition and scope issues*

We note at 5.6: “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.” This seems to undermine the purpose of the draft code as an over-broad carve out.

Digital Rights Watch is concerned that the focus on disinformation is limited to inauthentic behaviour, defined as, "spam and other forms of deceptive behaviours (including via automated systems) which encourages users of Digital Platforms to propagate content which may cause harm." It is not clear what is meant by ‘inauthentic’ and there is no clear

[1] Learn more about our work on our website: https://digitalrightswatch.org.au/

justification for using this term. We agree that fake accounts and bots pose a problem, and can see how this could be defined as ‘inauthentic behaviour.’ But the problem of disinformation is much larger than this. It also encompasses ‘authentic’ trolls, in other words, people who intentionally spread disinformation. Greater clarity about the terminology and justification for its use would be welcome.

The draft code states also that disinformation does not include "partisan news and commentary." It is not clear what constitutes partisan news, and further information about this definition would assist. Labelling something as "news” gives the impression it is being asserted and should be treated as fact.

It is unclear whether the scope of the draft code covers externally generated fake news that is shared by users (defined as ‘misinformation’ in the discussion paper). We have assumed that this would be considered “user generated content” at 4.1 but this should be made clear. Our suggestion is to change this wording to "user generated and/or user shared content." While the act of sharing externally generated fake news may not be accompanied by harmful intent, it is often part of a strategy deployed by those who do possess this intent and should be considered within the scope of the code.

We are concerned that it is not clear how paid advertising will be treated under the code. Paid advertising is not included in the scope of the code or the excluded services and products (at 4.1 and 4.2), however we also note Outcome 3 of the draft code. While the EU code is clear that misleading advertising is not considered disinformation, this seems to be a difficult position to maintain in a context in which sponsored posts from external sources are a common kind of disinformation. As set out on page 12 of the discussion paper, paid advertising is also part of the problem of disinformation because it used for financial gain.

It would be useful to know whether a partial or total ban on political advertising has been considered, or whether a partial and total ban on access to advertising by certain key websites that propagate disinformation has been considered. At the very least, we think that the code should encourage restrictions that are platform specific when the advertising buyer has engaged in spreading disinformation, eg restricting access to microtargeting, and require transparency reporting on these measures.

### *Accountability and transparency*

We note the Comment at 5.7 that “the risk that the release of certain information may result in an increase in behaviours that propagate Disinformation or which increase its virality.” We think transparency and accountability are essential for this code to be meaningful. Disingenuous and malicious users already game the rules of various platforms to spread disinformation. The risk has already materialised, and this should not be a reason to resist calls for accountability and transparency.

Users are an important resource for identifying disinformation in real time. For this reason, and noting Outcome 1c in the draft code, we are concerned that any complaints or reporting system is properly resourced. We believe it would be helpful to set benchmarks for response and resolution times (for example, a response within 24 hours and resolution within 48 hours other than in exceptional circumstances). Reporting on compliance with these benchmarks should become part of the annual reporting structure.

Digital Rights Watch is concerned that this code could inadvertently result in increased censorship of users, either because of its broad scope or because of the risk of policies being applied unevenly. If signatories are to implement and publish policies/guidelines on the prohibition and management of Disinformation (5.10), care should be taken to not disproportionately target particular groups over others, and to take steps to verify this is the case via reporting. For example, we have seen that Instagram's Community Guidelines have not been applied fairly and without discrimination in the past.

### *Research*

We note the Comment at 5.20: "Signatories commit not to prohibit or discourage good faith research into Disinformation on their platforms." The relationship between platforms and researchers has rarely been straight forward. We note that Facebook has appeared to do exactly what is discouraged by 5.20 in the case of the NYU Ad Observatory.[2] We think Objective 5 is critically important, and would encourage DIGI to work with its members and other stakeholders to find ways to give this greater meaning to this commitment in practice and ensure it is applied in a non-discriminatory way.

## Contact

**Lucie Krahulcova** | Programme Director | Digital Rights Watch | lucie@digitalrightswatch.org.au

[2] Facebook Seeks Shutdown of NYU Research Project Into Political Ad Targeting: https://www.wsj.com/articles/facebook-seeks-shutdown-of-nyu-research-project-into-political-ad-targeting-11603488533