European Commission public consultation on the European Democracy Action Plan

Mozilla submission annex

15 September, 2020

Mozilla is the Corporation behind the Firefox web browser, the Pocket "read-it-later" application and other products and services that collectively are used by hundreds of millions of individuals around the world. Mozilla's parent company is a not-for-profit Foundation that focuses on fuelling the movement for a healthy internet. Mozilla is also a global community of thousands of contributors and developers who work together to keep the internet open and accessible for all.

We support the European Commission's policy objectives for the European Democracy Action plan, and our response to the questionnaire aims to provide insight and guidance for the next stages of policy development. This annex supports our questionnaire response, and elevates a number of key policy considerations that we believe should be at the forefront of the Commission's reflections.

Importantly, our questionnaire responses should be read *in conjunction with* this annex, as unfortunately many parts of the questionnaire do not facilitate nuanced contextual responses.

1. The EU Code of Practice on Disinformation

Mozilla was an original participant in the European Commission High-Level Expert Group on Fake News and Disinformation, and we played a key role in the development of the EU Code of Practice on Disinformation in 2018.

1.1 Taking stock of how far we've come

At the outset it is important to acknowledge that the Code of Practice was a significant policy milestone. It was the first such instrument of its kind globally; it facilitated the exchange of information and best practice between public authorities and private companies on an urgent emerging policy challenge; and ultimately, it contributed to greater security and trust in the 2019 EU elections.

As a signatory to the Code of Practice we have sought to lead by example. In the past two years we have built tools within the Firefox browser to fight disinformation; empowered users with educational resources; supported research on the issue; and led advocacy efforts to push the other signatories to live up to their own commitments within the Code of Practice. Most recently, in context of our broader efforts under the Code we took steps to combat disinformation and promote authoritative information related to the COVID-19 pandemic, particularly through our Firefox Snippet and New Tab features, our Pocket curated media content service, and our Mozilla Foundation outreach and advocacy work. A complete overview of our COVID-19 efforts pursuant to the Code can be found here.

1.2 A Code fit for the future

Despite these achievements we have always been clear that, in policy terms, the Code is a *starting point.* There is considerably more work to be done, both to ensure that the Code's commitments are properly implemented, and to ensure that it is situated within a more coherent general EU policy approach to platform responsibility.

With respect to ensuring effective implementation of the Code's commitments, we note that in many instances implementation has been less successful than we had hoped. This is particularly evident with respect to signatories' commitments around advertising disclosure and empowering the research community. While these issues can be partially addressed in other legislative instruments (for instance, the Digital Services Act) they are key elements of the Code of Practice and should be considered as such. We therefore encourage the Commission to maintain a high-level of diligence in its monitoring of the implementation of signatories' commitments.

More broadly, we consider that the European Democracy Action Plan can provide an important opportunity to consider how the Code of Practice should fit within the broader EU approach to platform responsibility and content regulation. There are two pressing reasons why this is necessary -- first, the last EU political mandate coincided with a significant fragmentation in the policy approach to platforms and online content; and second, with the Digital Services Act the Commission is preparing a root and branch reform of the key principles of platform responsibility and content regulation.

We believe that the DSA and the Code of Practice can coincide in a coherent and effective regulatory paradigm. In our DSA public consultation submission we advance a vision of procedural accountability, whereby content responsibility should be assessed in terms of the Trust & Safety *processes* that platforms have in place to address illegal and harmful content on their services. Within a broad regulatory framework, platforms should be obliged to assess the various ways in which their services are at risk of illegal and harmful content, and to put in place commensurate Trust & Safety processes to address that risk. For instance, policy interventions could encourage enhancements to flagging systems or improvements to the means by which content is surfaced to users. This approach ensures interventions happen where they are likely to have the most impact in addressing and mitigating harm, but in a way that does not necessitate companies to unduly interfere with their users' fundamental rights.

The Code of Practice could serve as a concrete manifestation of this broader procedural accountability framework within the domain of disinformation. In theory under the Code, relevant companies assess the content-related problems they face, commit to taking certain steps to address them, and ultimately subject their efforts to oversight and assessment by the European Commission. To be a true manifestation of procedural accountability however, it is important that the efforts companies are making are *commensurate* to the risks they face, and that their interventions focus on practices and processes, rather than arbitrary 'outputs' (e.g. how much content was taken down in a given period of time). In that context, we again encourage the Commission to maintain diligence in assessing the implementation of commitments under the Code (to ensure they are commensurate) and to encourage companies to focus their efforts on practices and processes (e.g. providing better data access to researchers; providing better user-facing tools; etc).

Ultimately the Code of Practice has served an important role. Nonetheless there is still work to be done to maximise its potential and there is a pressing need to situate it firmly within a broader, thoughtful, platform accountability framework (preferably that which we advocate for in our DSA filing). The EDAP provides a timely opportunity to address both these matters.

2. Shining a light on disinformation through advertising disclosure

Today our ability to address disinformation as it manifests online is hindered by a lack of insight, and there are a number of areas where enhanced transparency could contribute to improved policy responses to disinformation. One area of particular importance is around ad placement, given that the paid-for advertising content has traditionally served as an important vector for the dissemination and amplification of disinformation.

One methodology to address this would be through the implementation of a framework whereby platforms that operate advertising networks publicly disclose all advertisements on their platforms via ad archive APIs.

If this approach was pursued, it could:

- Apply to all advertising, so as not to be constrained by arbitrary boundary definitions of 'political' or 'issue-based' advertising;
- Potentially include disclosure obligations that concern advertisers' targeting parameters for protected classes as well as aggregate audience demographics, where this makes sense given privacy and other considerations;
- Establish disclosure via publicly-available APIs, such that access is not restricted to specific privileged stakeholders as we have seen in some existing ads transparency efforts.

In the recent past, regulatory and co-regulatory initiatives aiming at ad transparency to combat disinformation have generally focused on 'political' advertising. Focusing on purely 'political' advertising (e.g. advertising copy developed by political parties) is too narrow an approach in many instances, and is often considered to be insufficient to capture the complex web of actors involved in politically-motivated disinformation online.

A broad ads disclosure framework could also drive transparency with respect to what is known as 'issue' advertising. Experience has shown how disclosure obligations that include this broader category of political ads put platforms in a challenging position, as the relevance for disclosure purposes of particular issue-based advertising requires platforms to decide what is 'political' in nature, which can vary depending on context, jurisdiction, and time. These problems could potentially be avoided by disclosure of *all* advertisements. It avoids the risk of under-disclosure (that arises with overly narrow definitions of ‘political’ advertising); minimises the burden on platforms to make highly-contextual definitional assessments (that arises with definitions of ‘issue-based’ advertising); and, helps ensure that transparency objectives are resilient in the face of technological and commercial changes (the nature of advertising content and channels evolves constantly).

Further, the inclusion of all ads allows for the identification and analysis of other forms of systemic harm that may be occuring in the current ad ecosystem. Indeed, other types of advertising that are not overtly political in nature may nonetheless be deceptive or may be targeted in a way that discriminates towards particular groups. For example, advertisements for jobs or housing may be targeted to certain demographic groups, in violation of fundamental rights.

A thoughtful analysis of how to balance privacy considerations, business considerations, and transparency is necessary for a successful transparency landscape. The inclusion of targeting parameters and aggregate audience demographics can be a significant tool for ensuring that regulators and researchers can understand *how* disinformation can spread across platforms. For instance, for much of the disinformation that is delivered via advertising on platforms, the content of the advertising provides only a partial - and indeed ancillary - insight into the phenomenon. Rather, it is the fact of *what types* of individuals those advertisements are aimed at and under *what circumstances*, that can provide insight into the risks and harms.

3. Microtargeting: Developing a meaningful problem definition

A number of the questions in the European Democracy Action Plan concern the microtargeting of online content, particularly content that is political in nature.

Related to section 2, we view highly sophisticated microtargeting and personalisation as an important contributor to the spread and impact of disinformation. Indeed, a key factor that allows disinformation to be impactful in the online ecosystem is that it can be targeted to those populations *most susceptible* to its messages. For instance, disinformation aimed at suppressing the vote for a specific candidate is most problematic when targeted at individuals *who intend to vote* for said candidate (see for instance, this research that concerns voter suppression disinformation aimed at African Americans in the US).

Yet despite this awareness, there is currently a dearth of well-developed policy options - such as ones that would establish the appropriate limits on targeting at some specified level of granularity - that could reasonably address the problem and which provide a clear articulation of the likely benefits and drawbacks of such approaches.

At the very least then, we should aim to establish a clear problem definition and evidence-base on which we can develop policy options. For that reason we suggest to consider including targeting parameters and aggregate audience demographics within the suggested framework for the bulk disclosure of all advertisements that we outline in section 2. We recognise that there can also be unintended consequences as well as business, security and privacy considerations at play here. And, thus, it will be important to navigate this potential option with stakeholders that can ensure that those considerations are part of the debate upfront about the benefits and drawbacks of this possible approach.

At the same time, it should also be underscored that the European Commission already possesses many legislative instruments that could be brought to bear to address harms arising from microtargeting. For instance, effective enforcement of the GDPR across the EU may be fruitful in allaying many of the concerns around collection and use of data for microtargeting of political content.

Ultimately, we welcome the Commission's consideration of the role of microtargeting with respect to political advertising and its contribution to the spread and impact of disinformation. The EDAP provides a crucial opportunity to systematically understand the problem and develop the necessary evidence base for effective policy responses.

4. Encrypted messaging apps: Addressing disinformation while maintaining trust and security

Recent experiences in the European Union during the COVID-19 pandemic have demonstrated that the inadvertent spread of both misinformation and disinformation on messaging apps is not a problem limited to emerging economies.

However, it must be underscored that the spread of disinformation via such platforms is not a consequence of their deployment of e2e encryption. On the contrary, these platforms may be vulnerable to disinformation owing to: gaps in digital literacy, unresponsive product design, and ineffective redress mechanisms. Keeping this context in mind, solutions should focus on these three categories, and not on weakening encryption. We urge the Commission to maintain its commitment, expressed in the questionnaire text, that measures to address disinformation should be developed "with full respect of encryption and data protection law".

4.1 Digital Literacy

The inadvertent spread of disinformation on e2e platforms is fundamentally a problem of digital literacy and a lack of awareness on how to verify news/information before forwarding it. Different stakeholders, including platforms, governments and media organisations, have a shared responsibility on this front. It is vital that e2e services provide users with sufficient resources on the importance and means of verifying news/information they receive. This could take the form of publicly available resources on websites and other media awareness campaigns.

4.2 Responsive Product Design

Despite many of the clearly recognised concerns with disinformation on e2e messaging services, many service providers have been slow in reforming their products to empower users with tools to combat disinformation while protecting e2e encryption. For example, while provided by some players in the industry, many services are yet to make it easier for users to be able to identify suspicious links and frequently forwarded messages in a user-friendly manner within their applications. As recent examples from other jurisdictions have demonstrated, it is possible to do so without breaking e2e encryption. Another measure that services could make available is allowing users to conveniently search the internet or fact-checking websites with the content of messages to verify their authenticity or be warned of possible harm that may occur from such information.

4.3 Redress Mechanisms

It is currently quite difficult for users to report messages and users to the e2e platform for action (including account suspension) despite overt actions which violate platform conditions and regulations. This should be reformed to enable users to report messages/users as easily as they can forward or delete them. To ensure complete transparency and similar to online backup feature in some services, this act by the user should explicitly state that doing so will mean the message will no longer be encrypted as it will (along with associated metadata) be shared with the platform for reporting and action. All of these measures can be in addition to the meaningful oversight, accountability, and appeal mechanisms mentioned in our response to the questionnaire.

****

In closing, we look forward to working alongside the Commission services to give practical meaning to the political ambition expressed in the European Democracy Action Plan, as well as the EU Code of Practice.

For further information on our consultation response and the positions expressed therein, please feel free to contact us at brussels@mozilla.com.