



**Monitoring of the COVID19 Disinformation – Summary of  
ERGA’s activities  
(ERGA Report)**

## **Table of contents**

<b>Executive Summary</b> .....	<b>3</b>
<b>Introduction</b> .....	<b>6</b>
<b>COVIDCHECK – Assessing the Implementation of EU Code of Practice on Disinformation</b> .....	<b>9</b>
<b>Cooperation with Signatories and Research Community focusing on Disinformation</b> .....	<b>13</b>
<b>Conclusions and Recommendations</b> .....	<b>19</b>

## Executive Summary

The EC's Joint Communication on 'Tackling COVID-19 disinformation – Getting the Facts right' resulted regarding the timeframe of the current study - between August 2020 and June 2021<sup>1</sup>- in the provision and publication of 47 monthly transparency reports issued by signatories of the Code of Practice. As noted in previous ERGA Reports, the Code of Practice is a unique, global example of an effort to establish cooperation between private companies hosting digital platforms, public authorities, non-governmental bodies, academia, media, and many other relevant stakeholders to fight disinformation. All these stakeholders have an important role to play in finding transparent solutions to the problem of misinformation and disinformation and ERGA appreciates the input and engagement of all these players since the Code was established. However, the pandemic has shown that massive circulation of disinformation online can cause serious harm to citizens across the EU and beyond. Research published by ERGA since the Code was introduced has highlighted significant weaknesses in the Code that need to be addressed if the Code of Practice is to become a more effective tool in fighting disinformation.

Commission Guidance on Strengthening the Code of Practice on Disinformation<sup>2</sup> issued in May 2021 reflected several of ERGA's recommendations included in reports published in 2019<sup>3</sup> and 2020<sup>4</sup>. ERGA endorses the Commission's expectations set out in the Guidance regarding a strengthened Code and believes that the outcomes of the most recent monitoring process presented here, further underpin the need for significant changes in the Code as set out in the Guidance. This includes increased transparency and greater accountability in terms of the implementation of the commitments in the Code by the signatories. ERGA is engaging with the other stakeholders in the process to strengthen the Code and welcomes the proposals in the Guidance for an enhanced monitoring framework that also includes representatives from ERGA and EDMO in a permanent Taskforce chaired by the Commission. ERGA believes that such a new forum is needed to ensure that the Code can be further evolved and issues related to the Code can be addressed in a more timely and collaborative manner. ERGA believes that the Taskforce should be established as soon as possible to progress discussions about the development and implementation of a suitable monitoring framework for the new Code.

---

<sup>1</sup> This is the timeframe for the considered period of this report. The monitoring programme on COVID-19 disinformation continues.

<sup>2</sup> Commission presents guidance to strengthen the Code of Practice on Disinformation, [https://ec.europa.eu/commission/presscorner/detail/en/ip\\_21\\_2585](https://ec.europa.eu/commission/presscorner/detail/en/ip_21_2585)

<sup>3</sup> Interim Assessment of the Implementation of the Code of Practice on Disinformation, [https://erga-online.eu/wp-content/uploads/2019/06/ERGA-2019-05\\_Statement\\_ERGA\\_work\\_on\\_disinformation-adopted1.pdf](https://erga-online.eu/wp-content/uploads/2019/06/ERGA-2019-05_Statement_ERGA_work_on_disinformation-adopted1.pdf)

<sup>4</sup> ERGA Report on Disinformation: Assessment of the implementation of the Code of Practice, <https://erga-online.eu/wp-content/uploads/2020/05/ERGA-2019-report-published-2020-LQ.pdf>

ERGA believes this report and its ten recommendations listed here should contribute to the strengthening of the Code of Practice as envisaged in the European Democracy Action Plan<sup>5</sup> (EDAP) and detailed further in the Guidance.

Recommendation 1: Improve the provision of more granular and country-specific data including samples of pieces of content as examples of enforced policies and information on the impact of activities.

Recommendation 2: Agree on a more unified or standardised structure of reports.

Recommendation 3: Set-up a publicly available up to date country-level database of trustworthy and authoritative sources as part of the transparency centre.

Recommendation 4: The Code of Practice should explicitly clarify what forms of content the signatories act upon.

Recommendation 5: Platforms need to set up a central repository for relevant policies within the transparency centre that contains all relevant information regarding all signatories in a centralised manner.

Recommendation 6: The provision of data on the use and performance of automated systems should be an explicit commitment within the Code of Practice.

Recommendation 7: Access to data for independent research should be binding under the Code of Practice.

Recommendation 8: The monitoring framework needs to include standardised procedures verifying the implementation of actions across Member States.

Recommendation 9: The monitoring framework should include a provision for co-operation between the platform signatories and relevant authorities.

Recommendation 10: The commitment by the Signatories to appoint and fund an independent auditor to report on their activities to implement the commitments in the Code should be strengthened.

ERGA notes that the majority of recommendations were also included in its 2020 Report and are addressed in the Commission's Guidance. The recommendations were presented at a stakeholder workshop on October 8<sup>th</sup> which included signatories, the EU Commission, EDMO, ERGA, researchers, academics and fact-checkers. An initial draft was also presented to the Assembly of the signatories of the Code of Practice that is drafting the revised Code where the

---

<sup>5</sup> European Democracy Action Plan: making EU democracies stronger, [https://ec.europa.eu/commission/presscorner/detail/en/ip\\_20\\_2250](https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2250)

report and its recommendations were also welcomed as an important input to the revision of the Code.

It is important to stress that in light of the [Digital Services Act](#)<sup>6</sup> (DSA) proposal, ERGA sees the strengthened Code as an opportunity to test some of the proposals in the DSA related to access to data, audits, external oversight, or risk-mitigating measures. This report should serve as the initial step in that direction. All of these elements are an inherent part of a co-regulatory approach called for by EDAP and by the EU Guidance, and, in the opinion of ERGA members, adopting this approach may be the only way to address the shortcomings of the Code of Practice in its current form.

---

<sup>6</sup> The Digital Services Act: ensuring a safe and accountable online environment, [https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment\\_en](https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment_en)

## Introduction

The global pandemic caused by the new coronavirus highlighted, among many other things, the urgency of finding effective ways to limit the spread and impact of misinformation and disinformation online. False and intentionally manipulative content related to the virus itself, or restrictions imposed by national governments, and most recently to vaccines causes real harm in the offline world. In extreme cases, health-related misinformation and disinformation may even lead to additional deaths. The European Union and its Member States reacted<sup>7</sup> in 2020 by calling for stronger and more coordinated action to counter efforts of actors spreading propaganda or hatred. Digital Services Act and the European Democracy Action Plan are the flagship initiatives in this area and the European Regulators Group for Audiovisual Media Services (ERGA)<sup>8</sup> as an advisory body to the Commission continues to provide the expertise of its members to fine-tune and effectively implement them. In May 2020 ERGA published its Report on Disinformation<sup>9</sup> assessing the implementation of the Code of Practice on Disinformation. The main recommendations of the report are still relevant and remain to be substantially addressed.

In June 2020, the European Commission issued a Joint Communication ‘Tackling COVID-19 disinformation – Getting the Facts right’<sup>10</sup> which established a COVID-19 monitoring programme for the six platforms that are signatories of the EU Code of Practice on Disinformation<sup>11</sup>. Signatories were asked to provide monthly transparency reports between September 2020 and June 2021 which would highlight how the platforms were tackling COVID-19 disinformation and these reports were published online<sup>12</sup> by the European Commission. This was the third monitoring programme the Commission had established, having previously carried out targeted monitoring during the European Elections in May 2019 and a first-year assessment<sup>13</sup> of the Code

---

<sup>7</sup> Tackling coronavirus disinformation, [https://ec.europa.eu/info/live-work-travel-eu/coronavirus-response/fighting-disinformation/tackling-coronavirus-disinformation\\_en](https://ec.europa.eu/info/live-work-travel-eu/coronavirus-response/fighting-disinformation/tackling-coronavirus-disinformation_en)

<sup>8</sup> The European Regulators Group for Audiovisual Media Services brings together heads or high level representatives of national independent regulatory bodies in the field of audiovisual services, to advise the Commission on the implementation of the EU’s Audiovisual Media Services Directive (AVMSD).

On 3 February 2014, the European Commission adopted a Decision on establishing the ERGA and setting the objectives for the Group: to advise and assist the Commission in its work, to ensure a consistent implementation of the AVMSD as well as in any other matters related to audiovisual media services within the Commission’s competence; to facilitate cooperation between the regulatory bodies in the EU, as provided for in the directive regulating audiovisual media services and to allow for an exchange of experience and good practices.

<sup>9</sup> ERGA, *ERGA Report on Disinformation, Assessment of the implementation of the Code of Practice*, April 2020 <https://erga-online.eu/wp-content/uploads/2020/05/ERGA-2019-report-published-2020-LQ.pdf>

<sup>10</sup> Tackling coronavirus disinformation, [https://ec.europa.eu/info/live-work-travel-eu/coronavirus-response/fighting-disinformation/tackling-coronavirus-disinformation\\_en](https://ec.europa.eu/info/live-work-travel-eu/coronavirus-response/fighting-disinformation/tackling-coronavirus-disinformation_en)

<sup>11</sup> Code of Practice on Disinformation, <https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation>

<sup>12</sup> First baseline reports – Fighting COVID-19 disinformation Monitoring Programme, <https://digital-strategy.ec.europa.eu/en/library/first-baseline-reports-fighting-covid-19-disinformation-monitoring-programme>

<sup>13</sup> Assessment of the Code of Practice on Disinformation – Achievements and areas for further improvement, <https://digital-strategy.ec.europa.eu/en/library/assessment-code-practice-disinformation-achievements-and-areas-further-improvement>

in September 2020. Similar to previous assessment programmes, the Commission requested ERGA with evaluating the platforms transparency reports.

As with the previous requests, ERGA established a monitoring group in its Sub-Group working on disinformation and media plurality. Due to limited resources (financial and staffing) and the lack of a clear national mandate in this area, not all ERGA members were in a position to participate in this process. The monitoring group included ERGA members from Croatia, France, Germany, Greece, Hungary, Italy, Ireland, Latvia, Slovakia. The approach adopted in each country was determined by a range of factors including available resources, available information and other relevant activities that were in place nationally to tackle disinformation. In addition, ERGA through the European Commission sent an additional data request to the signatories in April 2021 and held a series of meetings with the signatories to discuss the process. A key focus of this information request was country-specific data for the Member states participating in the monitoring process. Unfortunately, no useful information was provided by the signatories by the deadline.

In the majority of cases, the monitoring was carried out by the relevant ERGA members in-house or through discussions and collaborations with relevant national experts. In May 2021, ERGA issued an interim report which recognised that the platforms appeared to have increased their efforts to tackle COVID-19 disinformation. However, ERGA also noted that due to the lack of country-specific data, many of the actions reported by the platform signatories could not be sufficiently assessed.

In addition, following a tender process, the Irish regulator, the BAI<sup>14</sup>, commissioned researchers at the Institute for Future Media, Democracy and Society (FuJo<sup>15</sup>) at Dublin City University (an EDMO hub) to undertake a detailed monitoring review of all the transparency reports that had been provided by the platform signatories and published by the Commission. This involved a mainly quantitative analysis of the value of the reports at an EU level and some qualitative case studies that were focused on the Irish experience. The emerging report, CovidCheck<sup>16</sup>, was published on September 16<sup>th</sup> 2021. CovidCheck found that while the EU Code has proven a useful instrument in prompting signatories to respond to concerns about disinformation, there are shortcomings in relation to its implementation and scope. Researchers cited difficulties in assessing the timeliness, completeness and impact of the actions undertaken by the signatories. The report (included in Appendix One) sets out nine recommendations for more effective reporting and monitoring of disinformation. The findings and recommendations in the

---

<sup>14</sup> New report recommends development of robust procedures for reporting and monitoring online disinformation, <https://www.bai.ie/en/new-report-recommends-development-of-robust-procedures-for-reporting-and-monitoring-online-disinformation/>

<sup>15</sup> <https://fujomedia.eu/>

<sup>16</sup> New report recommends development of robust procedures for reporting and monitoring online disinformation, <https://www.bai.ie/en/new-report-recommends-development-of-robust-procedures-for-reporting-and-monitoring-online-disinformation/>

CovidCheck report were endorsed by the ERGA monitoring group as they reflected the general experience in participating Member States and the findings of their own monitoring processes.

In September 2021, the French regulator the CSA also provided information on its second annual assessment<sup>17</sup> of measures taken by platforms to tackle disinformation. This report is a statutory requirement of French law and assesses the measures taken to tackle COVID-19 disinformation based primarily on reports provided to the CSA by relevant platforms. In the report, the CSA welcomes the increasing quantity and quality of data provided by the platforms and their willingness to cooperate with the regulator. However, it also notes that some major and crucial points still require further collaboration and transparency from the stakeholders. The report issues 16 recommendations and these also helped to inform the ERGA group findings.

The ERGA recommendations in this report were presented at a stakeholder workshop on October 8<sup>th</sup> which included signatories, the EU Commission, EDMO, ERGA, researchers, academics and fact-checkers. An initial draft was also presented to the Assembly that is drafting the revised Code where the report and its recommendations were also welcomed as an important input to the revision of the Code.

---

<sup>17</sup> Lutte contre la manipulation de l'information : le CSA publie le bilan des mesures mises en œuvre par les plateformes en ligne en 2020, <https://www.csa.fr/Informer/Toutes-les-actualites/Actualites/Lutte-contre-la-manipulation-de-l-information-le-CSA-publie-le-bilan-des-mesures-mises-en-oeuvre-par-les-plateformes-en-ligne-en-2020>. The executive summary in English is available at the bottom of the dedicated page on the CSA's website.



## COVIDCHECK – Assessing the Implementation of EU Code of Practice on Disinformation

The report<sup>18</sup> commissioned by the Broadcasting Authority of Ireland prepared by the Institute for Future Media, Democracy and Society at the Dublin City University provides further systematic analysis of the transparency reports between August 2020 and April 2021, complementing the assessment prepared by the ERGA Subgroup.

The analysis of 47 reports performed by the DCU involved manual coding and an automated textual analysis providing both quantitative, as well as qualitative data.

The following section represents the critical analysis of the reports by the DCU confirming the findings of ERGA and providing additional ones, endorsed by ERGA:

### Analysis of Reported Actions by Dublin City University

#### Reported actions

Between August 2020 and April 2021, the signatories collectively submitted 47 reports. Each individual action referenced in the reports was coded as a discrete item, resulting in the identification of 1114 individual actions (see Table 2)<sup>19</sup>. The number of actions reported by the signatories varied greatly. However, direct comparisons are complicated by the different sizes and nature of the signatories' operations. For example, Mozilla is the only signatory that does not operate a social media platform and confined its updates to just two reports. In contrast, Google operates a wide range of platforms and applications including Ads, Images, Maps, News, Play, Search, and YouTube. Similarly, Facebook operates three large-scale platforms - Facebook, Instagram, and WhatsApp - as well as a large advertising business. Google reported the highest number of actions (387). TikTok reported the lowest number of actions (87) among the signatories that operate a social media platform. Unsurprisingly, the highest number of actions were reported in August 2020 when signatories were asked to provide baseline reports outlining all actions undertaken to counteract COVID-19 disinformation between the start of the pandemic and the end of July 2020.

---

<sup>18</sup> E. Culloty et al, *COVIDCHECK, Assessing the implementation of EU Code of Practice on disinformation in relation to COVID-19*, DCU Institute of Future Media, Democracy and Society, 16 September 2021, <https://www.bai.ie/en/news-and-information/publications/>.

<sup>19</sup> E. Culloty et al, *COVIDCHECK, Assessing the implementation of EU Code of Practice on disinformation in relation to COVID-19*, DCU Institute of Future Media, Democracy and Society, 16 September 2021, <https://www.bai.ie/en/news-and-information/publications/>, p. 17.

## **New and continuing actions**

There was considerable repetition in the reporting of actions across the time period. In some instances, previous actions were referenced to provide new updates or to report an extended application. In other instances, the same action was presented in successive reports without new information. To gain a better understanding of how many new actions were undertaken in response to the pandemic, each action was coded as a new action or a continuing action, beginning with the baseline reports (see Table 3)<sup>20</sup>. If an action was in place prior to the COVID-19 pandemic, such as a media literacy initiative that began in 2019, then it was classified as a continuing action when it was reported in the baseline report of August 2020. Of the 1114 reported actions, only 351 (32%) were new while 763 (68%) were repetitions of previously reported actions. Notably, the majority (58%) of actions reported by Facebook concerned new actions whereas only 12 percent of Google's reported actions concerned new initiatives. Broken down by month (see Table 4)<sup>21</sup>, the vast majority of actions presented in the baseline reports of August 2020 were new initiatives. In subsequent reports, Facebook, Twitter and, to a lesser extent, TikTok regularly presented new actions while Google and Microsoft largely presented continuing actions.

## **Action types**

While signatories were asked to report on all policies and actions to address COVID-19 disinformation, the following areas were highlighted by the Commission: initiatives to promote authoritative content at the EU and Member State level; initiatives and tools to improve users' awareness; manipulative behaviour on their services; data on flows of advertising linked to COVID-19 disinformation; and support for factchecking. A quarter of all actions concerned the promotion of authoritative content such as links to information by the WHO or national health authorities (see Table 5)<sup>22</sup>. The next most common action areas were advertising responses (17%) and blocking, removing or demoting content (13%). The emphasis placed on these action areas varied across the platforms (see Table 6)<sup>23</sup>. Promoting authoritative content was the most common type of action reported by each signatory, with the exception of Facebook. For Facebook, the most common action type related to blocking, removing, or demoting content, which accounted for more than a fifth of all Facebook actions. Almost a quarter of all actions reported by TikTok concerned factchecking and content labelling and one fifth related to blocking, removing, or demoting content. Mozilla did not have actions in a number of categories, reflecting the fact that it does not provide a social media type service. Apart from Mozilla, all signatories operate significant advertising businesses. They were asked to provide information on advertising linked to COVID-19 disinformation including policies to limit such adverts on their own services and the placement of such adverts on third-party websites. Google reported the largest percentage of actions relating to advertising, which accounted for 23 percent of all its actions.

---

<sup>20</sup> Ibid.

<sup>21</sup> Ibid.

<sup>22</sup> Ibid, p.19.

<sup>23</sup> Ibid.

## **Collaborations**

Just over a quarter of all actions were undertaken in collaboration with third-party organisations such as the WHO (see Table 7)<sup>24</sup>. Among the signatories, TikTok reported the highest level of collaborative actions at 46 percent followed by Twitter at 38 percent. Unsurprisingly, certain types of action attracted a higher level of collaboration including factchecking, the promotion of authoritative content, public health and media literacy, and research (see Table 8)<sup>25</sup>. These actions often involved working with factchecking partners, health agencies, and non-profit organisations. Almost half of all research actions were reported as collaborations although the overall number of research actions is relatively low. Notably, only four percent of actions relating to both content moderation (i.e. blocking, removing, demoting content) and organised manipulation were reported as collaborative endeavours.

## **Relevance to COVID-19**

Although signatories were asked to report on policies and actions that addressed COVID-19 disinformation, the reported actions by Facebook, Google, and Twitter were sometimes unrelated to the topic (see Table 9)<sup>26</sup>. For example, they reported on marketing workshops that provided support for start-ups and the launch of media literacy campaigns ahead of elections. Twitter and Facebook, in particular, dedicated large segments of their reports to describing charity work and general public-health initiatives. The links to COVID-19 were often tenuous. For example, Twitter reported the launch of a new campaign and emoji to commemorate International Holocaust Memorial Day, which was linked to COVID-19 by referencing the rise in hateful and racist rhetoric during the pandemic. It should be noted, however, that signatories were asked to report “all instances of social media manipulation, malign influence operations or coordinated inauthentic behaviour detected on their services”. This may be interpreted as an invitation to report instances of organised manipulation even when those instances were unrelated to COVID-19. When those cases were removed, Google’s percentage of irrelevant actions fell to nine percent, but the Facebook and Twitter percentages remained high at 28 percent each (see Table 10)<sup>27</sup>.

## **Regional application**

The Commission requested data relating to the EU and at a Member State level. The actions described in the report concerned a mix of regions (see Table 11)<sup>28</sup>. In many instances, it was difficult to discern which regions were covered by the reported actions as no specific region was mentioned or geographic reach was vaguely defined as “available in 32 countries”. In total, 13 percent of actions were referenced as having a global reach. Regarding the EU, 34 percent of

---

<sup>24</sup> Ibid.

<sup>25</sup> Ibid.

<sup>26</sup> Ibid, p. 21.

<sup>27</sup> Ibid, p. 21

<sup>28</sup> Ibid.

actions covered all of the EU while 12 percent covered some but not all EU Member States. The regional application was unclear or unstated for 40 percent of actions. It is possible that these cases were intended to be read as global actions. However, as new initiatives are frequently rolled out in a phased manner across markets, these actions were designated unstated or unclear in the absence of clarity in the reports. One percent of actions did not apply to any EU Member State. These mainly concerned the US market and were irrelevant to the scope of the Code. Across the signatories, Google and Microsoft reported the lowest number of actions where the regional application was unstated or unclear and the highest number of actions that were stated to be applicable to all EU Member States (see Table 12)<sup>29</sup>. Both Google and Microsoft offered clarity about the regional application of their reported actions. At the beginning of each report, Microsoft stated: “we generally track and report these efforts on a global or EU-wide basis. In those instances where our efforts are limited to a certain Member State, we have stated that below”. Similarly, Google frequently clarified that: “unless specified otherwise, the content of this [report] section applies equally to all EU Member States.” Consequently, it was generally possible to determine the regional application of actions for these signatories. In contrast, the regional application could not be determined for more than two-thirds of the actions reported by Twitter and Facebook and half of the actions reported by TikTok. Actions that applied only partially to the EU ranged from single-country actions, such as a media literacy program in France, to actions that applied to most, but not all, Member States. Most of the non-EU actions reported by Facebook and Twitter applied to the United States. Mozilla reported one action applying to Africa; this accounted for six percent of the total due to Mozilla’s small number of overall actions.

Research findings and recommendations of both the ERGA and the research institutions, including both university-affiliated and those not affiliated with any university, as well as with fact-checkers outlined in this report support and strengthen each other. ERGA believes this cooperation, such as for example with the Dublin City University represents a best practice and should be an integral part of the future monitoring framework. Strong and specific commitments by signatories, clear mandate for regulators within ERGA for monitoring of its implementation, combined with the access to data for independent researchers and their effective cooperation will substantially contribute to fulfilling the aim of creating an effective monitoring framework.

---

<sup>29</sup> Ibid.

## Cooperation with Signatories and Research Community focusing on Disinformation

As mentioned in the introduction to this report, ERGA members participating in the monitoring reached out to several research institutions, both academic and non-academic, and to individual researchers to provide further context. The main focus was put specifically on the data from February 2021 that would be useful in verifying the claims made by the signatories in their March 2021 reports. At the same time, ERGA organised a series of meetings with the signatories of the Code, namely Facebook, Microsoft, Twitter, Mozilla, TikTok, and Google. Request for additional data was circulated to the signatories and the bilateral meetings were scheduled to discuss the details.

### **ERGA request for additional data circulated to the signatories**

#### General request: Disaggregate the reported data

Where possible, please provide country-specific data (numbers on steps taken as well as policies announced) already highlighted in the transparency reports in each of the categories below.

#### 1) Source of authoritative content

What institutions have been identified by the platform as authoritative sources on COVID-19 related content in each of the monitored countries? Provide a country-specific list. What are the criteria used to qualify an institution as an “authoritative source”?

Provide user engagement metrics (views/clicks (click-throughs)/shares) of the content of authoritative sources for each monitored country for 1 – 28 February 2021.

#### 2) Measures to limit the spread of covid 19 disinformation

Please provide data for 1 – 28 February 2021

The number of pieces of content, related to COVID-19 and in particular COVID-19 vaccines, tagged or indexed as such on each online service provided. Provide data for each monitored country in the monitored period.

The number of pieces of content, related to COVID-19 and in particular Covid19 vaccines, removed and/or demoted from each online service provided. Provide data for each monitored country in the monitored period.

User engagement metrics - number of views/clicks/shares, that such pieces of content received before being removed and/or demoted. Provide data for each monitored country in the monitored period.

Country-specific sample of accounts and/or groups blocked, closed, or banned as a result of a violation of terms of services related to COVID-19 together with reasons for blocking, closing, or banning for each. Samples will remain for internal purposes only.

Please provide a representative sample of pieces of content posted by these accounts and/or groups that led to the measure implemented (e.g. blocking, banning, closing). Samples will remain for internal purposes only.

Provide country-specific statistics on how many (percentage) decisions were made by automatic systems and how many were made by human moderators.

### 3) Notice and action mechanism – flagged/fact-checked content

Please provide data for 1 – 28 February 2021

The number of pieces of content reported/flagged by users, percentage of flagged content acted upon (removed, demoted, labelled), and their respective number/percentage of the latter. Provide data for each monitored country.

Percentage of content related to COVID-19 and in particular vaccines that have been fact-checked and labelled in each of the monitored countries in the monitored period.

The number of user interactions with content related to COVID-19 and vaccines that have been fact-checked and labelled as containing false claims (before and after labelling). Provide data for each monitored country.

Please provide the country-specific representative samples for the monitored period for each of the categories above per monitored country. Samples are for internal purposes only.

### 4) For platforms with trusted flaggers:

Please provide data for 1 – 28 February 2021

Provide a list of institutions/organisations designated as trusted flaggers for your platform in each monitored country in relation to COVID-19 content.

Number of COVID-19 – related pieces of content flagged by trusted flaggers. Provide data for each monitored country.

### 5) Advertising

Please provide data for 1 – 28 February 2021

The number of ads related to COVID-19 and in particular vaccines labelled as such in each monitored country.

The number of ads rejected in each of the monitored countries containing disinformation related to COVID-19 and in particular vaccines. Such a number should be disaggregated from the number of ads rejected because classified as a scam.

#### 6) Additional data:

##### 1) Messaging apps

What policies and specific measures have been implemented to limit the dissemination of COVID-19 related disinformation in the messaging apps in monitored countries? Please provide country-specific quantitative data on actions taken.

All signatories provided additional context related to the granularity of data during the meetings and Facebook, TikTok and Twitter provided additional data in late April 2021.

Facebook provided additional data on cooperation on campaigns promoting the social distancing guidance or campaigns on the importance of washing hands and wearing a mask with the European Parliament and national public health authorities in Italy, Denmark, or Finland. In addition, Facebook set up a set of public CrowdTangle Live Displays allowing users to see both global and country-level data on COVID-19 related content.

Slovak Ministry of Health for example confirmed they received free advertising credits and they also observed a significant rise in reach and impact of their content on Facebook. According to representatives of the Health Ministry in Slovakia, Facebook also offered assistance in creating an advertising campaign.

TikTok reacted to ERGA request for additional data with an update on solutions related to mis/disinformation introduced already by the end of 2020, namely notice Tags on COVID-19 content, search banners leading to WHO or national authorities' websites, updated FAQ in the COVID-19 Center and Hashtag PSAs leading to authoritative sources. Given the level of cooperation with fact-checkers for example in the EU Member States, the quantitative data provided by TikTok cover only 4 countries: Italy, Germany, Spain, and France. According to TikTok, the markets in other EU countries do not see the level of activity above the threshold worth reporting. The data in other countries therefore fluctuate and its accuracy is, according to TikTok, questionable, therefore not reported.

Twitter also submitted their answers to the request for data from ERGA. Specifically, the answers include additional insights on the #GetTheFacts prompt on COVID-19 vaccines leading users to national and global public health authorities. A list of relevant authorities per the Member States was included as well. The report further claims the country-specific data on pieces of content acted on as part of COVID-19 misinformation policy was lacking at the time of the delivery to ERGA.

On a general level, granularity and the scope of data (specifically the country-specific data) submitted by the platforms was very limited but there is an agreement between signatories and ERGA to continue the conversation.

### **External Data from Researchers and Organisations**

The country-specific data provided by the platforms in response to the ERGA request at the start of April has been very limited so the country-specific monitoring had to rely heavily on information from academics, civil society and other interested groups.

Two specific examples in Slovakia<sup>30</sup> and Lithuania<sup>31</sup> hint at the fact that the Facebook's measure<sup>32</sup> had a limited effect in these Member States. The following examples also highlight the need for wider discussion on the interplay between signatories, monitoring framework (as envisaged in the strengthened Code of Practice), and relevant research institutions in the Member States.

EU DisinfoLab, an independent NGO focusing on research and tackling sophisticated disinformation campaigns has shared several reports related specifically to COVID-19 and US elections-related disinformation and responses of the online platforms. Their publication *One Year Onward: Platform Responses to COVID-19 and US Elections Disinformation in Review*<sup>33</sup> offers a detailed timeline of responses to the disinformation of Facebook, Google, Twitter, and TikTok, specifically. The report highlights the efforts of platforms to promote authoritative sources and acknowledges the novel policy of providing free advertising credits to public health authorities and significant grants for journalism and fact-checking. The timeline also records the policy updates of the monitored platforms.

Below is an illustrative example of the timeline database mapping the measures by platforms per response type:

---

<sup>30</sup> Gerulata Technologies, a Slovak company focusing on monitoring and analysis of disinformation reported that in February 2021, out of 45 posts containing COVID-19-related disinformation posted by Facebook pages associated with disinformation, none were removed at the time. In addition, posts of these pages when including disinformation were 4-times more viral than other posts on the same page. Interestingly, in March 2021 Facebook removed a page responsible for 16 (of the 45 cited above) posts from February with COVID-19 disinformation. The page re-appeared with a slightly changed new name and already gathered more than 1000 followers in a month.

<sup>31</sup> In Lithuania, DebunkeEU Disinformation Analysis Center reported several Facebook pages spreading COVID-19 disinformation on the platform while being administered from the same Facebook profile in Lithuania. Specifically, they highlighted two pages: COVID vakcinų diskusijos (Discussions on Covid vaccines) and COVID vakcinų atsiliepimai" (COVID vaccine reviews). The pages are still active on Facebook.

<sup>32</sup> Kang-Xing Jin, Keeping People Safe and Informed About the Coronavirus, <https://about.fb.com/news/2020/12/coronavirus/>

<sup>33</sup> Alaphilippe A., Mayer T., *One Year Onward: Platform Responses to COVID-19 and US Elections Disinformation in Review*, EU Disinfo Lab <https://www.disinfo.eu/publications/one-year-onward-platform-responses-to-covid-19-and-us-elections-disinformation-in-review/>



Disinformation events response (black panels, given COVID-19, see 18 above)	Platform	Content	Jan	Feb	March	April	May	June	July	Aug	Sept	Oct	Nov	
Flagging/labelling content	Facebook	provides reports from sharing organisations, increases use of AI to detect misinformation and deep fakes			commits to invest in additional funding for fact-checking, news and research on misinformation, content moderation and user home	provides educational program for users who reacted to harmful COVID-19 misinformation, content moderation partly return to office		starts labelling of state-controlled media	starts adding labels to posts on voting, including before of elections and candidates		launches second wave of labelling for fact-checking and news, includes on-page labels on posts from institutions and campaigns that prematurely claim victory or attempt to delegitimise the elections	provides resources campaigns of public health agencies on vaccines		
	Google	continues work with fact-checking organisations, signals that shared articles are checked (age and not subject to double-checking fact-checking)				provides 60% in funding from the Google News Initiative for fact-checking								
	TikTok	provides reports from sharing organisations		starts attaching a "trusted" flag to COVID-19 related with a message "Learn the facts about COVID-19 that respects the user to verify facts, trusted news or information"							incidents fact-checking partnerships to help verify election-related information, adds an in-app reporting option for election misinformation	links to elections guide at the bottom of videos relating to the election and on videos from verified political accounts	includes an in-app announcement on hashtag pages relating to US President Trump and COVID-19 to remind users to verify facts with authoritative sources, Public Community Guidelines, and report content they believe may violate our policies.	
	Twitter			introduces labels and warning messages on synthetic and manipulated media			introduces labels and warning messages on potentially harmful, misleading COVID-19 related content				expands Civil Integrity Policy to label or remove false or misleading information intended to undermine public confidence in an election or other civic process (previously published only about presidential control, not/for)	sets of election right labels Tweets making claims about election results		
Blocking/removing content	Facebook	removes posts for removal of misleading manipulated media	removes COVID-19 disinformation "with immediate effect" from Facebook and Instagram	removes COVID-19 disinformation from recommendations "unless posted by a credible health organisation"			removes disinformation content that incites or suggests voting					removes false content from recommendations, secures page view restriction policies	sets rules for post warning when combined with additional language	
	Google					publishes YouTube policy on removal COVID-19 medical misinformation			expands Harmful Health Claims Policy for publishers to include content that contains authoritative scientific consensus on health risks				updates YouTube's COVID-19 Medical Misinformation Policy, Google Search Medical Topics Policy to include vaccine claims that contradict consensus of WHO and public agencies (see below on Oct and Nov in elections) using content removal or content through enforcement of existing policies on targeted individuals and organisations, hate speech, harassment, incitement to violence, spam, manipulated media, political advertising	
	TikTok					updates already existing policies on removing misleading or harmful content, to additionally remove non-verifiable medical information and to tackle the spread of COVID-19 related conspiracy theories					adds a policy that prohibits synthetic or manipulated content that misleads users by diverting the truth or events in a way that could cause harm			
	Twitter	continues to remove demonstrably false or potentially misleading content that has the potential of leading harm		removes synthetic and manipulated media flags to users/harm		incidents policy definition of harm to include content that contradicts COVID-19 public health guidance			provides additional details on factors taken into account when removing potentially misleading COVID-19 related content		updates Civil Integrity policy to label or remove false or misleading information intended to undermine public confidence in an election or other civic process			

The EU Disinfo Lab database of responses of the 4 major platforms points to a need of establishing a repository for disinformation-related policies of platforms. Dynamic changes in these policies and the fact that these changes are often announced in various places (for example blog posts) make it difficult to effectively assess or monitor their implementation in a given period.

The issue of monitoring and effectively diminishing the harm of disinformation shared through messaging apps was also highlighted by several research organisations. GLOBSEC, a think-tank in Slovakia publishing regular reports on disinformation and its effect on society in both Slovakia and the region of Central Europe, has pointed towards a trend of voice messages being shared widely through various messaging apps including Facebook messenger. There were several instances of such voice messages being shared during March and April 2020, at the beginning of the pandemic in Slovakia, and re-appeared again in March 2021. The more recent problematic and widely shared message is a woman's voice claiming COVID-19 vaccines have terrible side effects on vaccinated doctors and nurses and this information is being kept secret from the public. Such content poses a challenge to researchers trying to track, monitor, or report it to the platform. GLOBSEC reported on several narratives being shared by Facebook pages with a relatively large following leading to the rise of vaccine hesitancy in several regions of the country.

Question of transparency related to content moderation generally and actions taken against content flagged by research organisations, in particular, was raised during several meetings with the research community, including the interview with professor Fabio Giglietto of the University of Urbino who published three reports describing three networks of coordinated Facebook pages and their deceiving strategies<sup>34</sup>.

<sup>34</sup> An Italian Network of Memes, News, Spam Pages, and Dietary Supplements, [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3743531](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3743531); Adapting to Mitigation Efforts: Evolving Strategies of Coordinated Link Sharing on Facebook, [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3775469](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3775469); and

Slovak organisation DigiQ specialising in monitoring and analysing hate speech concluded that almost 43% of reported potential illegal and abusive content did not receive any feedback from the digital platforms hosting it. ERGA has since the publication of its Report on disinformation<sup>35</sup> called for more and meaningful transparency in content moderation and policies against manipulative behaviour.

The report commissioned by the BAI<sup>36</sup> and prepared by the Dublin City University provides additional valuable case studies generally confirming findings of previous ERGA reports dedicated to the monitoring of disinformation.

The case study dedicated to Facebook shows that fact-checking is not particularly effective in groups. Moreover, there are obviously problems in recognising problematic information from the same source (through cropping, re-contextualisation, etc.), and this has already been discovered in other previous studies. Content posted by politicians is not fact-checked (not even afterwards). Users are only very rarely warned when joining groups in which Covid-19 disinformation has been spread. DCU study also notes that moderators should be given more competencies to pre-filter content and that Facebook acts as a gateway to other platforms with less strict moderation practices.

A case study focusing on TikTok confirmed that the platform sponsors authoritative content in searches especially in the area of Covid-19, but not when it comes to Vaccinations (#vaccines, #vaxx), there seems to be a blind spot here. The study goes on to conclude that labels were applied rather unsystematically to problematic content.

The above-mentioned examples show that the strengthened Code of Practice on Disinformation as well as other mechanisms related to both illegal and harmful content need to reflect this as well. In addition, when establishing a monitoring framework related specifically to the Code of Practice on disinformation, the Key performance indicators should focus on transparency as well. Transparent reporting on the impact of both disinformation and policies deployed against it will lead to more effective monitoring and that, in turn, will lead to more effective policies.

---

Coordinated Hateful Disinformation on Italian Politics and Social Issues, since 2017, [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3777263](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3777263).

<sup>35</sup> ERGA, *ERGA Report on Disinformation, Assessment of the implementation of the Code of Practice*, April 2020 <https://erga-online.eu/wp-content/uploads/2020/05/ERGA-2019-report-published-2020-LQ.pdf>

<sup>36</sup> E. Culloty et al, *COVIDCHECK, Assessing the implementation of EU Code of Practice on disinformation in relation to COVID-19*, DCU Institute of Future Media, Democracy and Society, 16 September 2021, <https://www.bai.ie/en/news-and-information/publications/>

## Conclusions and Recommendations

The agreement between the Commission and platforms on the monthly transparency reports published by signatories of the Code of Practice is a globally unique example of an effort to establish cooperation between public authorities, non-governmental bodies, academia, media, and many other relevant actors. Platforms are an integral part of finding transparent solutions to the problem of misinformation and disinformation. Therefore, ERGA very much appreciates the willingness of the major platforms to take and implement commitments themselves as they have done in the framework of the Code of Practice on Disinformation. At the same time, the current pandemic situation has shown that disinformation has the potential to cause immediate physical harm and that the commitments are not enough. Media regulation stands for the protection of freedom of expression – in order to fulfil this task, the corresponding binding rules are now urgently needed as a basis for regulation.

In addition, an effective monitoring system is needed. This report is aimed to assess whether the transparency reports fulfil the initial aim of the reporting and monitoring exercise as set out in the European Commission's Communication from June 2020 and to propose how the monitoring of the implementation as laid out in the EU Commission's Guidance on Strengthening the Code of Practice on Disinformation<sup>37</sup> from April 2021 should be set up. According to the communication, in the second phase, the platforms and industry associations should have made use of specific indicators to enable a monthly overview of the effectiveness and the impact of their policies in curbing COVID-19 related disinformation. This report highlights the need for additional specific indicators to include, should there be any effective measurement of the effectiveness and the impact of policies and measures implemented and of content moderation deployed.

The ERGA assessment and accompanying recommendations should also contribute to the strengthening of the Code of Practice as envisaged in the European Democracy Action Plan and detailed further in the EU Commission's Guidance. It is important to highlight that in light of the Digital Services Act (DSA) proposal, ERGA sees the strengthened Code as an opportunity to test some of the proposals in the DSA related to access to data, audits, external oversight, or risk-mitigating measures. This report should serve as the initial step in that direction.

Below are the key findings accompanied by ERGA recommendations, based on ERGA's experiences regarding the monitoring of the implementation of commitments under the Code of

---

<sup>37</sup> European Commission, COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS, European Commission Guidance on Strengthening the Code of Practice on Disinformation (COM(2021) 262)

Practice. These recommendations aim to improve and strengthen the Code and its monitoring framework.

### **Finding 1: Insufficient data**

ERGA reports published in 2019<sup>38</sup> and 2020<sup>39</sup> highlighted the need for more detailed and in particular country level datasets to support effective monitoring. This has also been highlighted by others who have assessed the impact of the Code including VVA in its report<sup>40</sup> for the Commission. While some additional country level data was provided in the COVID-19 Reports, it lacked detail and consistency. The transparency reports provided some information on policy updates, new activities and new tools relevant to countering COVID-19 disinformation. However, there was a lack of detailed data on engagement with or impact of these initiatives. Such information is required to enable effective monitoring and to improve transparency and public accountability.

**Recommendation 1: Improve the provision of more granular and country-specific data including samples of pieces of content as examples of enforced policies and information on the impact of activities.**

Detailed country-specific information allowing the NRAs to examine the implementation of the signatories' commitments on a Member State level is essential for effective monitoring. Furthermore, country-specific data is crucial to monitor the effectiveness of the Code throughout the EU, which will be particularly crucial once it becomes a co-regulatory instrument. Such information should be provided in relation to the implementation of policies and other activities countering disinformation and their impact. These datasets should include quantitative and qualitative information and include samples of content acted upon and details of the action taken to allow for verification. It should also contain the content that have been most relayed or that have had the biggest number of impressions. Clear parameters for such the reporting should be defined in the Code and should take account of the proposed requirements in Article 23(2) of the proposed Digital Services Act.

### **Finding 2: Lack of uniform structure of reports**

---

The executive summary in English is available at the bottom of the dedicated page on the CSA's website <https://www.csa.fr/Informer/Toutes-les-actualites/Actualites/Lutte-contre-la-manipulation-de-l-information-le-CSA-publie-le-bilan-des-mesures-mises-en-oeuvre-par-les-plateformes-en-ligne-en-2020> .

<sup>38</sup> Interim Assessment of the Implementation of the Code of Practice on Disinformation, [https://erga-online.eu/wp-content/uploads/2019/06/ERGA-2019-05\\_Statement\\_ERGA\\_work\\_on\\_disinformation-adopted1.pdf](https://erga-online.eu/wp-content/uploads/2019/06/ERGA-2019-05_Statement_ERGA_work_on_disinformation-adopted1.pdf)

<sup>39</sup> ERGA Report on Disinformation: Assessment of the implementation of the Code of Practice, <https://erga-online.eu/wp-content/uploads/2020/05/ERGA-2019-report-published-2020-LQ.pdf>

<sup>40</sup> Study for the assessment of the implementation of the Code of Practice on Disinformation, <https://digital-strategy.ec.europa.eu/en/library/study-assessment-implementation-code-practice-disinformation>

The Commission Communication set out four main areas for platforms to focus on in their reports, namely: 1) initiatives to promote authoritative content at EU and Member State level; 2) initiatives and tools to improve users' awareness; 3) manipulative behaviour and 4) data on flows of advertising linked to COVID-19 disinformation. While these headings are generally followed in the reports, different structures within the sections make comparisons between them almost impossible. In addition, the monitoring showed that the same initiative was reported in a series of reports. While some improvements were noted in later reports, such as the inclusion of an Executive Summary, overall, the lack of a consistent structure made a meaningful monitoring of the Code's effectiveness very difficult.

### **Recommendation 2: Agree on a more unified or standardised structure of reports**

A more unified or standardised structure of reports is required to allow effective monitoring. Any agreement on the structure might reflect the different nature of the signatories' services, but it is essential to find a common structure that allows for effective overview. Signatories, when designing the standard form for their reports, should focus on avoiding unnecessary repetition. Reports should clearly state the specific policy associated with the reported action and the relevance of the action to the Code of Practice. Signatories should make it clear whether the reported action is a new one. When designing a standardised structure of reports, signatories shall include specific information about the regional application and the application across the EU Member States and in different languages.

### **Finding 3: Need for a publicly available database of authoritative sources**

Transparency reports indicate that the signatories focused quite extensively on promoting content from authoritative sources using various strategies and tactics related to their respective services. In order to assess the impact on such activity, a current and updated country-level database of trustworthy and authoritative sources with the corresponding data on engagement with the content (e.g., views/clicks/shares) should be provided.

### **Recommendation 3: Set-up a publicly available up to date country-level database of trustworthy and authoritative sources as part of the transparency centre**

A centralised database where relevant signatories would publish information on trustworthy and authoritative sources whose content they promote as part of their awareness raising activities should be provided. Corresponding data on engagement with this content (e.g., views/clicks/shares) is also required for more effective verification of the claims in the reports. Such a database should be a part of the transparency centre as defined by the Commission's Guidance.

#### **Finding 4: Need for explicit definition of types and forms of content acted upon**

It is unclear from the transparency reports whether content reported by the signatories as acted upon includes comments, messages in the messaging applications, videos, pictures, sound recordings etc. There is a clear need for explicit definitions of all types and forms of content hosted on platforms that contribute to the spread of disinformation or other forms of harmful content.

#### **Recommendation 4: The Code of Practice should explicitly clarify what forms of content the signatories act upon**

Given the wide range of forms of harmful content that contribute to the spread of disinformation, a broader and clearer set of definitions of relevant content should be specified in the Code of Practice. These definitions should be used for future reporting including statistics on the various types and forms of content demoted, deleted or otherwise acted upon by the Signatories.

#### **Finding 5: No single repository of all relevant information on mis- or disinformation policies and activities is available**

While the transparency reports provide an overview of policies and activities related to countering disinformation, it is not clear if these relate specifically to COVID-19. It is also unclear in some instances whether the provided data is related to actions taken against specifically COVID-19 related disinformation or if the data is related to a broader policy issue.

#### **Recommendation 5: Platforms need to set up a central repository for relevant policies within the transparency centre that contains all relevant information regarding all signatories in a centralised manner.**

The Guidance asks Signatories to create and maintain a publicly accessible transparency centre featuring among other elements a dashboard displaying relevant KPIs and policies employed with geographical and language coverage. ERGA's experience and reports have demonstrated that an effective monitoring framework requires Signatories to maintain a repository that consolidates and makes generally accessible all their policies and activities to counter disinformation. Dynamic changes in these policies and the fact that these changes are announced in various places make it difficult, if not impossible to effectively assess or monitor their implementation. The availability of such a publicly available repository centralising the relevant policies of all signatories and their services would also make it possible to compare the actions taken at that very time. The information in the repository should be constantly updated to reflect the current situation.

#### **Finding 6: Code of practice does not include enough commitments for reporting on the use of automated systems to combat disinformation**

While the signatories report on actions, such as application of generic content labels, the role of automated systems and artificial intelligence in content moderation remains unclear. The monitoring framework should include information on how AI and other automated systems are being deployed to counter disinformation. Such reporting should also address the fact that AI tools are not equally advanced in all EU languages which impacts on how these tools can be effectively deployed across regions and Member States.

**Recommendation 6: The provision of data on the use and performance of automated systems should be an explicit commitment within the Code of Practice**

Commitments under the Code should include provision of data on the use and performance of automated systems and Artificial Intelligence in particular. Signatories should report on their use of these systems against disinformation including explanations on languages covered and on the types of disinformation these systems are trained to detect as well as provide further technical information if required by the competent NRA. The reports should also include information on risk assessment conducted by the Signatories on these systems.

**Finding 7: Access to data for national regulatory authorities and independent researchers supporting the Commission and ERGA in monitoring activities is an essential condition for an effective monitoring framework.**

ERGA's Monitoring reports in 2019 and 2020 highlighted the necessity for access to data related to disinformation for the independent researchers. The Digital Services Act introduces obligations for platforms to make data available to "vetted researchers" (Art. 31) and this provision is echoed in the Guidance (8.1). While the existing initiatives of several platforms are welcome, it should not be solely up to platforms to decide what type of data is provided or the format in which the data is presented.

**Recommendation 7: Access to data for independent research should be binding under the Code of Practice**

The Code should ensure access to data on disinformation for independent researchers. Signatories should also commit to expanding and improving their services that allow researchers to access data (such as Crowdtangle in case of Facebook). Researchers from non-university-affiliated institutions and organisations e.g., Factcheckers should also be included. Verification of the validity of data delivered to researchers, e.g., through an independent audit should be an integral part of the Code.

**Finding 8: Actions are not applied consistently and an insufficient monitoring framework**

ERGA's monitoring activities confirm that the reported actions are not being applied consistently and the reporting on the engagement metrics needs to be enhanced. Analysis of the self-

assessment and the ERGA experience also show there is a significant gap in the monitoring of the Code as it assesses what actions the signatories have reported without the certainty that those actions have been implemented across EU Member States and are working as stated. In addition, the Case Studies in the Irish report, indicated that there is a lack of consistency of applying labels on content containing disinformation and misinformation and that a more consistent approach is required. Verification of the reported actions is included as a requirement in the current draft of the Digital Services Act..

**Recommendation 8: The monitoring framework needs to include standardised procedures verifying the implementation of actions across Member States**

The strengthened Code needs a permanent monitoring framework that remains open to future changes. Such a framework, including its procedural aspects should be explicitly detailed in the Code of Practice and progressed by the Taskforce envisaged in the Commission Guidance. In addition, as outlined in the Guidance, in order for the new monitoring framework to be effective, it should be built on quantitative and qualitative KPIs capable of measuring the effectiveness of the commitments at an operational and structural level. The Code of Practice should include standardised procedures to verify the implementation of actions for future monitoring. Such procedures, if followed rigorously will ensure consistency in monitoring and provide an important counterpoint to the signatories' reported metrics. ERGA conclusions and experiences with monitoring of the Code in relation to both the European elections and the COVID-19 disinformation prove that without such procedures, the monitoring is very difficult and, in many instances, even impossible. As laid out in the Guidance, providers of online services that are widely used at EU level and have higher risk profiles with respect to the spread of disinformation should provide their reports within the monitoring framework every six months, while other signatories shall report on a yearly basis.

**Recommendation 9: The monitoring framework should include a provision for co-operation between the platform signatories and relevant authorities**

More intensive cooperation of signatories with relevant authorities and among the signatories themselves should be considered especially in critical periods such as during electoral campaigns or referendums in each Member State. In order to facilitate such of cooperation, signatories should designate a point of contact. The cooperation should furthermore also be developed on the proper technical expert level, e.g. trust and safety teams, where relevant. Furthermore, the monitoring framework's KPIs should include specific requirements in relation to reporting on key areas including content labels, content and account removals, fact-checking and media literacy campaigns. Relevant data should be made available through the transparency centre and via other adequate channels to the independent researchers for verification and to the other platform signatories.



**Recommendation 10: The commitment by the Signatories to appoint and fund an independent auditor to report on their activities to implement the commitments in the Code should be strengthened**

The existing Code includes a commitment to appoint a third-party organisation to review the annual self-assessment reports submitted by the relevant Signatories, and to evaluate the level of progress made against the Commitments. This has not been implemented to date. The revised Code should include this commitment and make provision for greater transparency about how this process will be implemented by each of the signatories. The Code should also include a commitment that the signatories will provide adequate resource for this process to be implemented annually. Such audit is crucial for the trust in the Code and the relevant data, and thus for an effective and trustworthy monitoring of the Code.

## Appendix

The CovidCheck Report can be accessed at the link below.

<https://www.bai.ie/en/new-report-recommends-development-of-robust-procedures-for-reporting-and-monitoring-online-disinformation/>

A video presentation of the findings by the researchers can be accessed at the link below.

<https://vimeo.com/606660750>