CODE OF PRACTICE ON DISINFORMATION

INTERMEDIATE TARGETED MONITORING – JANUARY REPORTS

On 29 January 2019, the Commission published the First Implementation Reports of the signatories to the Code of Practice on Disinformation, which detail policies and actions taken as of year-end 2018 to implement the full range of commitments the signatories have respectively subscribed to under the Code.

As called for in the Action Plan against Disinformation, the Commission is now carrying out targeted Monthly Intermediate Monitoring of the platform signatories' actions to implement Code commitments with particular pertinence to the integrity of elections, namely: scrutiny of ad placements (Commitment 1); political and issue-based advertising (Commitments 2 to 4); and integrity of services (Commitments 5 & 6).

The Commission has received Monthly Reports from **Google**, **Facebook** and **Twitter** addressing actions taken during January 2019 towards implementation of the commitments on electoral integrity. Broadly, the Commission is encouraged that the Reports provide further information on the policies the platforms have developed to meet these commitments. Nevertheless, the Commission remains deeply concerned by the failure of the platforms to identify metrics that would enable the tracking and measurement of progress in the EU as well as by lack of sufficient detail on the platforms' plans to ensure that actions in pursuit of their policies are being deployed in timely fashion and with appropriate resources across all Member States. **Google** has reported on actions taken during January to improve **scrutiny of ad placements** in the EU. These included 46,642 actions against EU-based Google Ads advertiser accounts for violation of the company's policies on misrepresentation and 3,258 actions against such accounts for violation of its policies on insufficient original content. As regards publishers, Google provided data on 205 actions against EU-based Ad Sense publishers for violation of the company's policies on valuable inventory; it did not detect any violations in January by EU-based Ad Sense publishers of the company's policies on misrepresentative content. While the Commission appreciates that the January Report offers per-Member State data, it considers that the metrics supplied do not elucidate the extent to which the actions were taken to address disinformation.

As regards **political and issues-based advertising**, Google published its new policy for "election ads" on 29 January; it is available in 25 EU languages. The policy covers ads that feature a political party, elected official or candidate for the European Parliament. Advertisers seeking to run such ads must be verified and document that the they are an EU-based entity or citizen of a Member State. Election ads will feature in-ad disclosure of the sponsor and will be included in Google's *Political Ads Transparency Report*, which Google expects to start publishing as soon as advertisers begin to run such ads. The Commission regrets that the January Report does not mark further progress with regard to the transparency and public disclosure of issue-based advertising.

As regards **integrity of services**, Google reports that it is staffing dedicated elections teams to prevent election-related abuse of its services, clamp down on malicious behaviour and react to breaking threats. It does not, however, provide detail on staffing or confirming coverage by Member

State or language. Regrettably, the January Report does not include metrics (aggregate or per-Member State) on enforcement of its policies to ensure the integrity of its services.

In addition, Google reports on a number of other activities carried out in January. These include new trainings for journalists on digital tools to strengthen research and reporting skills, trainings for election professionals and journalists focused on security risks and Google tools that address them, media literacy efforts, and an initiative on automated speaker verification aimed at advancing research and developing countermeasures against "deep fakes."

Facebook has not reported on activities undertaken in January with respect to **scrutiny of ad placements**. Regrettably, the report does not include any information or data about Facebook's efforts to disrupt advertising and monetisation incentives for purveyors of disinformation. One metric that could be provided is the number of pages and websites that were not allowed to run ads or use monetisation features such as Instant Articles because they were identified as regularly sharing false news. Information on the amount of content labelled "Severe" under the brand safety policy might also be instructive. The Report notes that Facebook has fact-checking partners in eight Member States covering seven official EU languages and that it plans to expand fact-checking partnerships with pan-European news agencies to cover more EU countries. Fact-checked information is used to demote false content, thereby demonetising purveyors of disinformation. It would be instructive for Facebook to provide indicators relating to these activities such as, for example, instances where the distribution of false or sensationalist content was reduced thanks to the changes in the News Feed ranking, or the number of items identified as false news, either following feedback from users or by third-party fact-checkers.

With regard to **political and issue-based advertising**, Facebook's pan-EU archive for political and issue advertising will be available in March 2019. No further information is provided regarding the features of this archive. We note, however, that on 11 February Facebook announced that a public API for machine search of the archive would be made available at the end of March. Facebook has also developed guidance for its political ads authorisation process, which will prevent political ads from running until advertisers complete the authorisation process and which will apply a "paid for by" disclaimer to such ads. The Report does not detail whether Facebook has started implementing this authorisation process in all Member States. Performance indicators for the archive will only become available when the service launches. The Report gives some general details regarding Facebook's classification of issue-based ads. Under its elections integrity programme, it will establish an elections operation centre in Dublin. Presumably, this centre would cover all Member States. However, no details are provided on the linguistic and/or other competences being deployed to cover the whole EU, nor is the centre's launch date indicated. Facebook continued its training activities for electoral stakeholders to safely and securely manage political communications on the platform. It continued partnering with electoral bodies and other relevant government bodies and agencies on electoral integrity issues, though the Report does not provide details on the number of Member States in which this cooperation is active.

On **integrity of services**, the Report provides an update on coordinated inauthentic behaviour, highlighting new cases of interference from third countries (Iran and Russia) that pertain to EU Member States (France, Germany and Romania, Latvia, Estonia, Lithuania,). However, the Report

does not provide information for January on the number of fake accounts removed due to malicious activities targeting specifically the EU.

The Report also describes Facebook's activities pertaining to other areas of the Code of Practice, such as media literacy and training for journalists.

Twitter reaffirmed its commitment to improving **the scrutiny of ad placements**, but it did not report any additional efforts in January or provide any metrics with respect to its commitments in this area. The Commission deeply regrets this omission and expects it to be addressed in Twitter's subsequent monthly reporting.

The January Report indicates that Twitter continues to work on **political ads transparency** for the European elections. It notes that, following the launch of Twitter's *Ads Transparency Center* and its work on the U.S. midterm elections, the company has been able to provide comprehensive data on paid election communications on its service, including purchases made by accounts, past and current ads served for accounts, targeting criteria and results, number of views received, and certain billing information associated with the accounts. The Report contains some aggregate metrics from the U.S. experience, but provides no specific information regarding the plans for rolling out its transparency tools in the EU. The Commission takes note of Twitter's announcement of 19 February that it will start enforcing its policy on political campaigning ads on 11 March 2019. However, it had expected Twitter to report such details in January, as promised in Twitter's First Implementation Report, and urges Twitter to provide them in February, including a confirmation that its political ads transparency tool will be available in all Member States before the start of campaigning for the European elections. The Commission also regrets that the January Report does not report on progress with regard to the transparency and public disclosure of **issue-based advertising**.

As regards **integrity of services**, the January Report highlights the addition of five new account sets, comprising numerous accounts in Russia, Iran, Bangladesh, and Venezuela, to its *Archive of Potential Foreign Operations*, which are publicly available and searchable. However, while the Report reaffirms Twitter's commitment to strengthen its services against manipulation through malicious automated accounts, spam, and other activities in violation of the company's terms of services, the Commission is concerned by the lack of metrics enabling the measuring of progress in this area.

Twitter also reported on a new research initiative focused on improving the performance of Machine Learning in social systems, such as Twitter.