

Code of Practice on
Disinformation – Report of
[AIForensics] for the period
[March 2023–March 2024]

Table of Content

Executive summary	3
V. Empowering Users	1
Commitment 17.....	1
Commitment 28.....	2
Commitment 29.....	2
VIII. Transparency Centre	3
Commitment 34.....	3
IX. Permanent Task-Force	3
Commitment 37.....	3
X. Monitoring of Code	4
Commitment 38.....	4
Commitment 39.....	4
Commitment 40.....	4
Commitment 42.....	5
Reporting on the service's response during an election	6
European Elections.....	7

Executive summary

Executive summary (max. 2 pages)

As the Code matures and Signatories are learning to work together under a common framework, AI Forensics' engagement remains centred on our organisation's two key areas of focus: - algorithmics auditing and participation in the subgroups on Generative AI and the Elections Monitoring Working Groups of the Crisis Response Subgroup. With the Code of Practice transitioning to a Code of Conduct in 2024, we are looking forward to keep engaging with other Signatories, the Commission, ERGA and EDMO.

V. Empowering Users

Commitment 17

In light of the European Commission's initiatives in the area of media literacy, including the new Digital Education Action Plan, Relevant Signatories commit to continue and strengthen their efforts in the area of media literacy and critical thinking, also with the aim to include vulnerable groups. [change wording if adapted]

Measure 17.1	Relevant Signatories will design and implement or continue to maintain tools to improve media literacy and critical thinking, for instance by empowering users with context on the content visible on services or with guidance on how to evaluate online content.			
QRE 17.1.1 [insert wording if adapted]	AI Forensics contributes to improve media literacy and critical thinking by publishing its work and disseminating it among media , stake-holders and decision makers. Furthermore, AI Forensics participates in conferences , meetings and events that serve as platforms to inform the larger public on the importance of algorithmic auditing, accountability and transparency.			
SLI 17.1.1 - actions enforcing policies above [change wording if adapted]	Number of conferences and events participated; number of media coverages			
	Total count of the tool's impressions	Interactions/ engagement with the tool	Other relevant metrics	Other relevant metrics
Data				
Measure 17.2	Relevant Signatories will develop, promote and/or support or continue to run activities to improve media literacy and critical thinking such as campaigns to raise awareness about Disinformation, as well as the TTPs that are being used by malicious actors, among the general public across the European Union, also considering the involvement of vulnerable communities.			
QRE 17.2.1 [insert wording if adapted]	AI Forensics is dedicated to increasing critical thinking among users and helping them restore their self-agency. Our innovative data-driven methodology provides journalists, researchers, and policymakers with timely evidence of systematic violations of users' interests and digital rights, particularly for minority groups and communities that are often overlooked in the design of technology. We believe that consistent and coordinated scrutiny is the path to restoring the balance of power between big tech platforms and its users. Therefore, AI Forensics will continue producing research and investigations that are fulfilling this aim.			
SLI 17.2.1 - actions enforcing policies above [change wording if adapted]	Methodology of data measurement [suggested character limit: 500 characters]			
	Nr of media literacy/ awareness raising activities organised/ participated in	Reach of campaigns	Nr of participants	Nr of interactions with online assets
Data				

Measure 17.3	For both of the above Measures, and in order to build on the expertise of media literacy experts in the design, implementation, and impact measurement of tools, relevant Signatories will partner or consult with media literacy experts in the EU, including for instance the Commission’s Media Literacy Expert Group, ERGA’s Media Literacy Action Group, EDMO, its country specific branches, or relevant Member State universities or organisations that have relevant expertise.
QRE 17.3.1 [insert wording if adapted]	

VI. Empowering the research community	
Commitment 28	
Relevant Signatories commit to support good faith research into Disinformation that involves their services. [change wording if adapted]	
Measure 28.1	Relevant Signatories will ensure they have the appropriate human resources in place in order to facilitate research, and should set-up and maintain an open dialogue with researchers to keep track of the types of data that are likely to be in demand for research and to help researchers find relevant contact points in their organisations.
QRE 28.1.1 [insert wording if adapted]	AI Forensics has in its team itself six researchers with multidisciplinary socio-technical background. Our team participates in open sessions and workshops at the universities, like, for example Winter and Summer Schools at the Amsterdam Universities. Furthermore, we have an inherently collaborative approach, and work in synergy with other research organisations in the ecosystem. Last year we collaborated with Algorithm Watch on investigating elections integrity in Bavarian, Hesse and Swiss elections . We released a report in collaboration with Check First “ An Analysis of Amazon’s Bookstore Algorithms within the DSA Framework ”. Furthermore, we have started research with the Stiftung Neue Verantwortung for the analysis of TikTok’s systemic risks under the DSA, its report will be published around April this year.
Measure 28.2	Relevant Signatories will be transparent on the data types they currently make available to researchers across Europe.
QRE 28.2.1 [insert wording if adapted]	AI Forensics releases their investigations and its methodology openly in their reports.
Measure 28.3	Relevant Signatories will not prohibit or discourage genuinely and demonstratively public interest good faith research into Disinformation on their platforms, and will not take adversarial action against researcher users or accounts that undertake or participate in good-faith research into Disinformation.
QRE 28.3.1 [insert wording if adapted]	AI Forensics works towards the goal of promoting good faith research.

VI. Empowering the research community	
Commitment 29	

Relevant Signatories commit to conduct research based on transparent methodology and ethical standards, as well as to share datasets, research findings and methodologies with relevant audiences. [change wording if adapted]	
Measure 29.1	Relevant Signatories will use transparent methodologies and ethical standards to conduct research activities that track and analyse influence operations, and the spread of Disinformation. They will share datasets, research findings and methodologies with members of the Task-force including EDMO, ERGA, and other Signatories and ultimately with the broader public
QRE 29.1.1 [insert wording if adapted]	All the research produced by AI Forensics is publicly available for anyone to access.
QRE 29.1.2 [insert wording if adapted]	
Data	

VIII. Transparency Centre	
Commitment 34	
To ensure transparency and accountability around the implementation of this Code, Relevant Signatories commit to set up and maintain a publicly available common Transparency Centre website. [change wording if adapted]	
Measure 34.1	
Measure 34.2	
Measure 34.3	
Measure 34.4	
Measure 34.5	

IX. Permanent Task-Force	
Commitment 37	
Signatories commit to participate in the permanent Task-force. The Task-force includes the Signatories of the Code and representatives from EDMO and ERGA. It is chaired by the European Commission, and includes representatives of the European External Action Service (EEAS). The Task-force can also invite relevant experts as observers to support its work. Decisions of the Task-force are made by consensus. [change wording if adapted]	
Measure 37.1	

Measure 37.2	
Measure 37.3	
Measure 37.4	
Measure 37.5	
Measure 37.6	
QRE 37.6.1 [insert wording if adapted]	AI Forensics is an active participant in the Monitoring and Reporting Subgroup, in the AI Generative Subgroup as well as in the Crisis Response Subgroup (where we participate in both in the Elections Steering Committee).

X. Monitoring of Code	
Commitment 38	
The Signatories commit to dedicate adequate financial and human resources and put in place appropriate internal processes to ensure the implementation of their commitments under the Code. [change wording if adapted]	
Measure 38.1	
QRE 38.1.1 [insert wording if adapted]	AI Forensics has two representatives, directly involved in the work of the Subgroup and Working group; we are part of ensuring full compliance with relevant Commitments taken under the Code.

X. Monitoring of Code	
Commitment 39	
Signatories commit to provide to the European Commission, within 1 month after the end of the implementation period (6 months after this Code's signature) the baseline reports as set out in the Preamble. [change wording if adapted]	

X. Monitoring of Code	
Commitment 40	
Signatories commit to provide regular reporting on Service Level Indicators (SLIs) and Qualitative Reporting Elements (QREs). The reports and data provided should allow for a thorough assessment of the extent of the implementation of the Code's Commitments and Measures by each Signatory, service and at Member State level. [change wording if adapted]	
Measure 40.1	

Measure 40.2	
Measure 40.3	
Measure 40.4	
Measure 40.5	
Measure 40.6	

X. Monitoring of Code

Commitment 42

Relevant Signatories commit to provide, in special situations like elections or crisis, upon request of the European Commission, proportionate and appropriate information and data, including ad-hoc specific reports and specific chapters within the regular monitoring, in accordance with the rapid response system established by the Taskforce. [change wording if adapted]

Reporting on the service's response during an election

Reporting on the service's response during an election

European Elections

Threats observed or anticipated at time of reporting: There is a fast growing influence of search and recommendation algorithms and search engines on electoral processes and increasing societal fractures and institutional distrust. Observing and identifying potential interferences and malpractices of the information ecosystem during elections dynamics provides a lot of essential information for evidence collection but also helps ensure the elections are fair and transparent.

Mitigations in place – or planned – at time of reporting: AI Forensics will run a cross-platform and cross-countries elections integrity investigation during the EU elections.

[Note: Signatories are requested to provide information relevant to their particular response to the threats and challenges they observed on their service(s). They ensure that the information below provides an accurate and complete report of their relevant actions. As operational responses to crisis/election situations can vary from service to service, an absence of information should not be considered a priori a shortfall in the way a particular service has responded. Impact metrics are accurate to the best of signatories' abilities to measure them].

Policies and Terms and Conditions

Outline any changes to your policies

Policy	Changes (such as newly introduced policies, edits, adaptation in scope or implementation)	Rationale

Scrutiny of Ads Placements

Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.

Specific Action applied (with reference to the Code's relevant Commitment and Measure)	Description of intervention

	Indication of impact (at beginning of action: expected impact) including relevant metrics when available
Specific Action applied (with reference to the Code's relevant Commitment and Measure)	Description of intervention
	Indication of impact (at beginning of action: expected impact) including relevant metrics when available
Political Advertising	
Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.	
Specific Action applied (with reference to the Code's relevant Commitment and Measure)	We are conducting research on the use and reach of political advertising during the current electoral period. Our results will be made available publicly, and should support enforcement monitoring of the commitments taken in this group.
	Distribution and press coverage of our findings, acceptance of our mitigation measure demands. .
Integrity of Services	
Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.	
Specific Action applied (with reference to the Code's relevant Commitment and Measure)	Description of intervention

	Indication of impact (at beginning of action: expected impact) including relevant metrics when available
Empowering Users	
Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.	
Specific Action applied (with reference to the Code's relevant Commitment and Measure)	We will run an investigation to evaluate the influence of algorithm-driven content dissemination during the EU elections. This comprehensive, cross-national, and cross-platform inquiry aims to analyze content from YouTube, TikTok, and Microsoft Copilot pertaining to election-related searches. AI Forensics will collect the platforms' responses and results, consisting of suggested videos for Youtube and TikTok and textual answers with links to sources for Copilot . The collected data will then be distributed among partners for manual categorization, focusing on the accuracy of information, the credibility of sources, and the detection of generative AI-produced content, whether audio or visual.
	Algorithmic auditing provides a solution on how algorithms can be held accountable and provide users with self-determination while navigating online. Auditing recommendation algorithms is critical to guarantee knowledge access on the internet since the new gatekeepers compromise that access. Algorithmic audits identify biases and malpractices and prevent exacerbation of inequalities and lack of self-agenda in knowledge consumption.
Empowering the Research Community	
Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.	
Specific Action applied (with reference to the Code's relevant Commitment and Measure)	We will run an investigation to evaluate the influence of algorithm-driven content dissemination during the EU elections. This comprehensive, cross-national, and cross-platform inquiry aims to analyze content from YouTube, TikTok, and Microsoft Copilot pertaining to election-related searches. AI Forensics will collect the platforms' responses and results, consisting of suggested videos for Youtube and TikTok and textual answers with links to sources for Copilot . The collected data will then be distributed among partners for manual categorization, focusing on the accuracy of information, the credibility of sources, and the detection of generative AI-produced content, whether audio or visual.
	Our investigation will increase the level of algorithmic accountability by lowering the barrier to entry for researchers to participate in the algorithmic audit space by collaboration in our project as partners. Collaborating partners will apply their expertise specific to each country to enhance the quality of the work by reviewing and refining the collection of prompts and queries, as well as labelling the gathered data. These partners will also take the lead in pre-election communication and outreach efforts targeting local media outlets. Academic collaborators will be responsible for publishing scholarly findings.

Empowering the Fact-Checking Community

Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.

Specific Action applied (with reference to the Code's relevant Commitment and Measure)

We will run an investigation to evaluate the influence of algorithm-driven content dissemination during the EU elections. This comprehensive, cross-national, and cross-platform inquiry aims to analyze content from YouTube, TikTok, and Microsoft Copilot pertaining to election-related searches. AI Forensics will collect the platforms' responses and results, consisting of suggested videos for Youtube and TikTok and textual answers with links to sources for Copilot . The collected data will then be distributed among partners for manual categorization, focusing on the accuracy of information, the credibility of sources, and the detection of generative AI-produced content, whether audio or visual.

Collaborating partners will apply their expertise specific to each country to enhance the quality of the work by reviewing and refining the collection of prompts and queries, as well as labelling the gathered data, those focusing their efforts on the fact-checking.