Code of Practice on Disinformation – Report of Adobe

Executive summary

Adobe has been a proud Signatory of the EU Code of Practice on Disinformation since June 2022, and we support the intention and ambition of this Code.

Adobe is a global leader in digital marketing and digital media solutions. Since the company's foundation in December 1982, we have pushed the boundaries of creativity with products and services that allow our customers to create, deploy, and enhance digital experiences. Our purpose is to serve the creator and respect the consumer, and our heritage is built on providing trustworthy and innovative solutions to our customers and communities. Adobe has a long history of pioneering innovation: when Adobe thinks about AI, we balance innovation with responsible innovation.

We are witnessing extraordinary challenges to trust in digital content. As social platforms amplify the reach and influence of certain content, mis-attributed and mis-contextualized content spreads quickly. Whether inadvertent misinformation or deliberate deception, inauthentic content is on the rise.

With the increasing volume and velocity of digital content creation, including synthetic media, it is critical to ensure transparency and restore trust in what we are consuming online. Adobe feels a responsibility to support the creative community and society at large and is committed to finding technical solutions that address the issues of manipulated media and tackle disinformation.

Content provenance and media literacy are a major focus for Adobe and the work of the Content Authenticity Initiative (CAI), which Adobe co-founded in 2019. We are focused on cross-industry participation, with an open, extensible approach for providing transparency for digital content (i.e. images, audio, video, documents, and generative AI) to allow for better evaluation of that content.

The Content Authenticity Initiative (CAI) now has over 3,300 members working to increase trust in digital content through provenance tools, which are the facts about the origins of a piece of digital content. The CAI works in tandem with the Coalition for Content Provenance and Authenticity (C2PA), an open technical standards organization also co-founded by Adobe in 2021, to implement our solution to combating misinformation online – called Content Credentials. Content Credentials are essentially a "nutrition label" for digital content – showing when a piece of content is created and modified, including whether AI is used. Content Credentials are a combination of cryptographic metadata and watermarking, designed to remain securely attached and to travel with the digital content wherever it goes. They include important information which may include the creator's name, the date an image was created, what tools were used to create an image and any edits that were made along the way. This empowers users to create a digital chain of trust and authenticity. The CAI developed free, open-source tools based on the C2PA standard for anyone to implement Content Credentials into their own products, services, or platforms.

This year, several major developments concerning the C2PA took place, with Google, Meta and OpenAI joining the C2PA steering committee and showing support of the C2PA standard and Content Credentials. TikTok also joined the C2PA as a member and began labeling AI-generated content uploaded to its platform with Content Credentials.

The Adobe-led CAI has also invested in creating and promoting media literacy curricula to educate the public about the dangers of deepfakes, the need for skepticism, and tools available today to help them understand what is true. In partnership with the Adobe Education team, the CAI updated their media literacy curriculum in February 2024 to include Generative AI curricular materials.

This open standard is more important than ever as powerful technology like Generative AI makes it easier to create, scale, and alter digital content and our work will continue to evolve and address the latest trends and landscape needs. We see Adobe’s focus on supporting and promoting wide adoption of content provenance tools as being particularly relevant to the EU Code of Practice on Disinformation and are grateful that Commitments relating to provenance and the C2PA open standard have been adopted as commitments in the Code in the Empowering Users chapter. We encourage all relevant Signatories to sign up to these commitments and join this cross-industry effort to tackle disinformation through technology.

Guidelines for filling out the report

Crisis and elections reporting template

Relevant signatories are asked to provide proportionate and appropriate information and data during a period of crisis and during an election. Reporting is a part of a special chapter at the end of the harmonised reporting template and should follow the guidelines:

- The reporting of signatories' actions should be as specific to the particular crisis or election reported on as possible. To this extent, the rows on "Specific Action[s]" should be filled in with actions that are either put in place specifically for a particular event (for example a media literacy campaign on disinformation related to the Ukraine war, an information panel for the European elections), or to explain in more detail how an action that forms part of the service's general approach to implementing the Code is implemented in the specific context of the crisis or election reported on (for example, what types of narratives in a particular election/crisis would fall into scope of a particular policy of the service, what forms of advertising are ineligible).
- Signatories who are not offering very large online platform services and who follow the invitation to report on their specific actions for a particular election or crisis may adapt the reporting template as follows:
    - They may remove the "Policies and Terms and Conditions" section of the template, or use it to report on any important changes in their internal rules applicable to a particular election or crisis (for example, a change in editorial guidelines for fact-checkers specific to the particular election or crisis)
    - They may remove any Chapter Section of the Reporting Template (Scrutiny of Ads Placement, Political Advertising, Integrity of Services etc.) that is not relevant to their activities
- The harmonised reporting template should be filled in by adding additional rows for each item reported on. This means that rather than combined/bulk reporting such as "Depending on severity of violation, we demote or remove content based on policies X, Y, Z", there should be individual rows stating for example "Under Policy X, content is demoted or removed based on severity", "Under Policy Y, content […]" etc.
- The rows should be colour-coded to indicate which service is being reported on, using the same colour code as for the overall harmonised reporting template.

Reporting should be brief and to the point, with a suggested character limit entry of 2000 characters.

Uploading data to the Transparency Centre

The reports should be submitted to the Commission in the form of the pdf via e-mail to the address CNECT COP TASK FORCE CNECT-COP-TASK-FORCE@ec.europa.eu within the agreed deadline. Signatories will upload all data from the harmonised reporting template to the Transparency Centre, allowing easy data access and filtering within the agreed deadline. It is the responsibility of the signatories to ensure that the uploading takes place and is executed on time. Signatories are also responsible to ensure that the Transparency Centre is operational and functional by the time of the reports' submission that the data from the reports are uploaded and made accessible in the Transparency Centre within the above deadline, and that users are able to read, search, filer and download data as needed in a user-friendly way and format.

Reporting on the signatory’s response during an election

Reporting on the signatory’s response during an election

2024 European Parliament Elections

Threats observed at time of reporting: AI-generated or AI-manipulated audio, video, and images that deceptively fake or alter the appearance, voice, or actions of political candidates, election officials, and other key stakeholders in a democratic election, or that provide false information to voters about when, where, and how they can vote.

Mitigations in place during the electoral period: [suggested character limit: 2000 characters].

[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	Description of intervention

relevant Commitment and Measure)	Indication of 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 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)	Description of intervention
	Indication of impact including relevant metrics when available
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	Description of intervention

relevant Commitment and Measure)	Indication of 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)	At the Munich Security Conference in February 2024, Adobe together with 19 other leading technology companies, pledged to help prevent deceptive AI content from interfering with this year's global elections. The "Tech Accord to Combat Deceptive Use of AI in 2024 Elections" is a set of commitments to deploy technology countering harmful AI-generated content meant to deceive voters. Signatories pledge to work collaboratively on tools to detect and address online distribution of such AI content, drive educational campaigns, and provide transparency, among other concrete steps. It also includes a broad set of principles, including the importance of tracking the origin of deceptive election-related content and the need to raise public awareness about the problem. Digital content addressed by the accord consists of AI-generated audio, video, and images that deceptively fake or alter the appearance, voice, or actions of political candidates, election officials, and other key stakeholders in a democratic election, or that provide false information to voters about when, where, and how they can vote. "Attaching provenance signals to identify the origin of content where appropriate and technically feasible" is one of the accord´s principle goals and the signatories commit to taking the following steps through 2024 with regard to provenance: Developing and implementing technology to mitigate risks related to Deceptive AI Election content by: a. Supporting the development of technological innovations to mitigate risks arising from Deceptive AI Election Content by identifying realistic AI-generated images and/or certifying the authenticity of content and its origin, with the understanding that all such solutions have limitations. This work could include but is not limited to developing classifiers or robust provenance methods like watermarking or signed metadata (e.g. the standard developed by C2PA or SynthID watermarking). b. Continuing to invest in advancing new provenance technology innovations for audio video, and images. c. Working toward attaching machine-readable information, as appropriate, to realistic AI-generated audio, video, and image content that is generated by users with models in scope of this accord.
	Indication of impact including relevant metrics when available

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)	Description of intervention
	Indication of impact including relevant metrics when available
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)	Description of intervention
	Indication of impact including relevant metrics when available