Code of Practice on Disinformation – Report of Adobe for the period 1 January – 31 December 2024

Table of Content

Executive summary	3
II. Scrutiny of Ad Placements	1
Commitment 1	1
Commitment 2	5
Commitment 3	6
IV. Integrity of Services	7
Commitment 15	7
V. Empowering Users	8
Commitment 17	8
Commitment 20	9
IX. Permanent Task-Force	11
Commitment 37	11
X. Monitoring of Code	12
Commitment 38	12
Reporting on the service's response during an election	13

Executive summary

Since June 2022, Adobe has been a Signatory of the EU Code of Practice on Disinformation, supporting the intention and ambition of this Code. In September 2024, Adobe also announced our commitment to uphold the pledges of the EU AI Pact, reiterating our support for the transparency requirements under the EU AI Act.

Adobe is a global leader in digital marketing and digital media solutions. Since the company's founding 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. As we continue harnessing the power of AI, it is critical to pair innovation with responsible innovation in order to ensure this technology is developed and deployed in a way that benefits everyone.

We stand at a pivotal moment as AI has begun to transform the way we live, work and play. As AI becomes more prevalent however, we are also witnessing extraordinary challenges to trust in digital content. In today's digital world, misattributed and miscontextualised content spreads quickly. Whether inadvertent misinformation or deliberate deception, inauthentic content is on the rise and once we are fooled once, we will begin to doubt anything we see or hear online – even if it is true.

With the increasing volume and velocity of digital content creation, including synthetic media, it is critical to ensure transparency in what we are consuming online. Adobe is committed to leading in this space and finding technical solutions that address the issues of manipulated media and deceptive digital content.

Content provenance and media literacy are a major focus for Adobe and the work of the <u>Content Authenticity Initiative (CAI)</u>, which Adobe co-founded in 2019 and leads today. 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 AI) to allow for better evaluation of that content.

The Content Authenticity Initiative (CAI) now has more than 4,000 members globally 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 organisation also co-founded by Adobe in 2021, to implement C2PA's solution to digital content provenance – called Content Credentials.

Content Credentials are essentially a "nutrition label" for digital content that anyone can implement to show how a piece of content is created and modified. Content Credentials are a combination of cryptographic metadata, fingerprinting and watermarking, designed to remain securely attached and travel with the digital content wherever it goes (for more information, please see "durable content credentials"). 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, including whether AI was used. 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.

In 2024, some major developments concerning the C2PA took place, with several companies joining the C2PA steering committee and demonstrating support for Content Credentials. TikTok also joined the C2PA as a member and began labelling Al-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 scepticism, and tools available today to help them understand what is true. In partnership with the Adobe Education team, the CAI updated its media literacy curriculum in February 2024 to include Generative AI curricular materials.

Provenance solutions such as Content Credentials are more important than ever as generative AI makes it easier to create, scale, and alter digital content. As AI rapidly evolves, our work will continue to adapt to emerging trends and evolving industry needs. We see Adobe's focus on supporting and promoting wide adoption of Content Credentials as being particularly relevant to the EU Code of Practice on Disinformation and are encouraged 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.

II. Scrutiny of Ad Placements Commitment 1 nit to defund the dissemination of disinformation, and imported for monetisation and ad placement, and the data to refer to the commitment of the data to refer to the data to the data

Relevant signatories participating in ad placements commit to defund the dissemination of disinformation, and improve the policies and systems which determine the eligibility of content to be monetised, the controls for monetisation and ad placement, and the data to report on the accuracy and effectiveness of controls and services around ad placements. [change wording if adapted]

of controls and serv	ices around ad placements. [change	wording if adapted]		
Measure 1.1	[insert wording if adapted]			
QRE 1.1.1 [insert wording if adapted]	Outline relevant actions [suggested cha	-		
	misinformation and disinformation. Itaken are the following: (1) Research was done to locate so Disinformation Index, CheckMyAds at (2) Flagged Sites were reviewed and Politifact, and MediaBiasFactCheck. (3) Domains where misinformation of (4) Historical impression reports were (5) Incidents found during the third (6) Adobe Advertising Cloud has read dis/misinformation. No new services	verified through manual checks of 30 or disinformation was confirmed were be pulled to assess the impression deliviperiod of submission have been added the control of the consumple of the consu	ing political content to be distrib disinformation by referencing rd party verification services suc added to the Service's Global Blay very on the domains. If to the tracker. Illation on available services with	outed via Adobe's Services. Actions g 3rd party reports from Global h as Global Disinformation Index, ocklist. h relevant solutions to combatting
SLI 1.1.1 –	Methodology of data measurement [su	ggested character limit: 500 characters]		
Numbers by actions enforcing policies above	Type of Action 1 [linked to the policy mentioned in QRE] 15 domains have been added to	Type of Action 2 [linked to the policy mentioned in QRE]	Type of Action 3 [linked to the policy mentioned in QRE]	Type of Action 4 [linked to the policy mentioned in QRE]
(specify if at page and/or domain level) [change wording if adapted]	Adobe's platform blocklist. This blocklist affects all transparent open market advertising.			
Level	Page/Domain	Page/Domain	Page/Domain	Page/Domain

Data		
Member States		
[example, insert only		
if feasible]		
Austria		
Belgium		
Bulgaria		
Croatia		
Cyprus		
Czech Republic		
Denmark		
Estonia		
Finland		
France		
Germany		
Greece		
Hungary		
Ireland		
Italy		
Latvia		
Lithuania		
Luxembourg		
Malta		
Netherlands		
Poland		
Portugal		
Romania		
Slovakia		
Slovenia		
Spain		
Sweden		
Iceland		
Liechtenstein		
Norway		
Total EU		
Total EEA		

1.1.1). It is based on		ries (query/bid ¹ or impression ²) and		onetise disinformation sources (under SLI nversion factor provided by a third party
SLI 1.1.2 -	Methodology of data measurement	[suggested character limit: 500 charac	cters]	
Preventing the				
flow of legitimate				ads verified as having dis/misinformation
advertising investment to sites	and using Ediquity's conversion ra	ate calculation to determine final va	liue in Euros.	
or content that are	Euro value of ads demonetised			
designated as	Early value of day demonetised			
disinformation				
[change wording if				
adapted]				
Data				
Measure 1.2	[insert wording if adapted]			
QRE 1.2.1 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters] Adobe Advertising Cloud does not permit political content. Adobe Advertising Cloud is a member of The Interactive Advertising Bureau (IAB), IAB Tech Lab, Network Advertising Initiative (NAI), Digital Advertising Alliance (DAA), European Interactive Digital Advertising Alliance (EDAA), and Trustworthy and Accountability Group (TAG).			
SLI 1.2.1 [change	Methodology of data measurement [suggested character limit: 500 characters]			
wording if adapted]	Nr of policy reviews	Nr of update to policies	Nr of accounts barred	Nr of domains barred
	No policies have been updated, added, or removed	None at this time	None at this time	15
Data				
Measure 1.3	[insert wording if adapted]			
	1. Advertisers are able to target/b	lock site domains or apps at the ca	mpaign placement level.	

¹ Request placed between a seller and buyer of advertising that can detail amongst other things website, specific content, targeting data inclusive of audience or content.

² Comprehensive calculation of the number of people who have been reached by a piece of media content by passive exposure (viewing a piece of content) or active engagement (visiting a destination).

	2. Advertisers are provided reporting on standard delivery metrics, primarily impression delivery, at a site domain and app level. 3. Advertisers are automatically opted into Adobe Ad Cloud's "Global Blocklist" which includes reviewed sites and apps determined to violate policies or are determined to be inappropriate for advertising. This prevents ad delivery on those properties (1) unless the advertiser has manually opted out or (2) the advertiser has entered into a private deal exposed to or not transparent with these properties. Sites and Apps are reviewed for brand safety, invalid traffic, and ad placement.
QRE 1.3.1 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters]
Measure 1.4	[insert wording if adapted]
QRE 1.4.1 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters]
Measure 1.5	[insert wording if adapted]
QRE 1.5.1 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters]
QRE 1.5.2 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters]
Measure 1.5	Outline relevant actions [suggested character limit: 2000 characters]
QRE 1.5.1 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters]
QRE 1.5.2 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters]
Measure 1.6	[insert wording if adapted]
QRE 1.6.1 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters] Adobe Ad Cloud offers several 3rd Party brand-safety targeting services that can be applied to campaign placements through our partners, with a fee. Pre-bid services halt impression delivery at the app, site or page level. These services are optional.
QRE 1.6.2 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters]
QRE 1.6.3 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters]
QRE 1.6.4 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters]
SLI 1.6.1 [change	Methodology of data measurement [suggested character limit: 500 characters]
wording if adapted]	In view of steps taken to integrate brand safety tools: % of advertising/ media investment protected by such tools:

Data		
Commitment 2		
Relevant Signato	ries participating in advertising commit to prevent the misuse of advertising systems to disseminate Disinformation in the form of	
Manage 2.1	advertising messages. [change wording if adapted] [insert wording if adapted]	
Measure 2.1 QRE 2.1.1 [insert wordi		
adapted]	Summe relevant detions [suggested character minu. 2000 characters]	
	Adobe Advertising Cloud's Ad Requirements Policy clearly states that false or misleading ads are prohibited. Additionally, political content is prohibited.	
	https://experienceleague.adobe.com/docs/advertising-cloud/policies/ad-requirements-policy.html?lang=en	
	Methodology of data measurement [suggested character limit: 500 characters]	
	If an ad is found or reported to be in violation of the Ad Requirements Policy, the ad placement is paused, the advertiser is notified to remove the ad and review Adobe Advertising's Ad Requirements Policy. If three separate event violations are found from the same advertiser, they will be removed from the platform.	
SLI 2.1.1 – Numbers by actions enforcing police above [change wording	policy mentioned in QRE] policy mentioned in QRE] policy mentioned in QRE] policy mentioned in QRE]	
adapted]	No advertisers have been found to violate Adobe Advertising Cloud's Ad Requirements Policy regarding disinformation or misinformation	
Data		
Measure 2.2	[insert wording if adapted]	
QRE 2.2.1 [insert wording adapted]	Outline relevant actions [suggested character limit: 2000 characters] Adobe Advertising Cloud is assessing services available from current and new partners for disinformation or misinformation. This includes block lists, measurement or reporting, and pre-bid services.	

Measure 2.3	[insert wording if adapted]	
QRE 2.3.1 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters] Ads are scanned by 3rd party partners to determine category. Ads from categories that pertain to Adobe Advertising Cloud's Ad Requirements Policy are flagged and reviewed for material that violates Adobe Advertising Cloud's Ad Requirements Policy.	
Methodology of data measurement [suggested character limit: 500 character Number of ads removed (as well as reach of ads before they were Num		aracters] Number of ads prohibited
SLI 2.3.1 [change wording if	successfully removed)	Number of aus prombited
adapted]	No ads were removed for violations of Adobe Advertising Cloud's advertising policy related to disinformation or misinformation.	
Data		
Measure 2.4	[insert wording if adapted]	
QRE 2.4.1 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters] If an advertiser is found to violate Adobe's Ads Requirements Policy, they are notified via email. This notice includes the specific ad in violation, the policy it violates, and a link to Adobe's Ads Requirements Policy. They are also notified that three violations may result in removal from the platform. Advertisers can appeal this process and be subject to a 90 day grace period to be reviewed if any additional violations are levied against them. The strikes may be removed after these conditions are met.	
	Methodology of data measurement [suggested character limit: 500 characters]	
SLI 2.4.1 [change wording if adapted]	Number of appeals 0	Proportion of appeals that led to a change of the initial decision
Data		

Commitment 3

Relevant Signatories involved in buying, selling and placing digital advertising commit to exchange best practices and strengthen cooperation with relevant players, expanding to organisations active in the online monetisation value chain, such as online e-payment services, e-commerce platforms and relevant crowd-funding/donation systems, with the aim to increase the effectiveness of scrutiny of ad placements on their own services. [change wording if adapted]

Measure 3.1	[insert wording if adapted]
ivieasure 5. i	[msert weranig in adapted]

QRE 3.1.1 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters]
	Adobe Advertising Cloud has partnerships with various 3rd party brand safety solution providers to offer targeting and reporting. Some have available targeting related to disinformation and misinformation. Adobe Advertising Cloud cooperates with supply partnerships to block any apps, sites, or sellers that are found to have disinformation and misinformation. Adobe Ad Cloud also reviews publicly available reports through reputable journals or news articles establishing specific acts of disinformation or misinformation.
Measure 3.2	[insert wording if adapted]
	Outline relevant actions [suggested character limit: 2000 characters]
QRE 3.2.1 [insert wording if adapted]	Adobe Advertising Cloud has had discussions with multiple 3rd party Brand Safety and Cybersecurity providers, exploring capabilities to target away from or block sites that host disinformation or misinformation.
Measure 3.3	[insert wording if adapted]
	Outline relevant actions [suggested character limit: 2000 characters]
QRE 3.3.1 [insert wording if adapted]	Adobe Advertising Cloud has integrated at least 15 domains found on publicly available reports to demonetise distributors of dis/misinformation.

IV. Integrity of Services

Commitment 15

Relevant Signatories that develop or operate AI systems and that disseminate AI-generated and manipulated content through their services (e.g. deep fakes) commit to take into consideration the transparency obligations and the list of manipulative practices prohibited under the proposal for Artificial Intelligence Act. [change wording if adapted]

Measure 15.1	[insert wording if adapted]
QRE 15.1.1 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters]

	Adobe supported the development and launch of Content Credentials, which act like a nutrition label for digital content and is fast becoming an industry standard for digital content provenance. The user-directed Content Credentials allow creators to provide attribution for their work. In the context of generative AI, Content Credentials can indicate whether a digital file was human-created, AI-edited, or AI-generated, allowing viewers to decide for themselves whether to trust it. Adobe automatically attaches Content Credentials to wholly generated Adobe Firefly outputs to indicate that generative AI was used to create it. This information enhances transparency for their audience.
Measure 15.2	[insert wording if adapted]
QRE 15.2.1 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters]

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]

adapted		
Measure 17.1	[insert wording if adapted]	
QRE 17.1.1 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters]	
	https://contentauthenticity.org/media-literacy	
	The Adobe-led CAI has also invested in creating and promoting media literacy curricula to educate the public about the dangers of deepfakes and tools available today to help them determine what's true online. In partnership with the Adobe Education team, the CAI updated our media literacy curriculum in February 2024 to include Generative AI curricular materials.	
	These standards-aligned lessons introduce students to generative AI and engage them in critical conversations surrounding the technology	
	Methodology of data measurement [suggested character limit: 500 characters]	

SLI 17.1.1 - actions enforcing policies above [change wording if adapted]	Total count of the to impressions	ool's	Interactions/ engagement v tool	vith the	Other rel	evant metrics	Oth	er relevant metrics
Data								
Measure 17.2	[insert wording if ad	lapted	i]					
QRE 17.2.1 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters]							
SLI 17.2.1 - actions enforcing policies above [change								
wording if adapted]	Nr of media	Rea	ch of	Nr of participants		Nr of interactions		Nr of participants
3 1 2	literacy/	cam	paigns			with online ass	ets	(etc)
	awareness raising							
	activities							
	organised/							
	participated in							
Data								
Measure 17.3	[insert wording if adapted]							
QRE 17.3.1 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters]							
Relevant Signatories commit to empower users with tool wording if adapted]	Commit s to assess the prove			tory or au	thenticity	or accuracy of	f digit	al content. [change
Measure 20.1	[insert wording if adap	ted]						
	C2PA standard and C	onten	t Credentials					
QRE 20.1.1 [insert wording if adapted]	Outline relevant action	s [sug	gested characte	er limit: 200	0 character	s]		
	Adobe is a co-founder and steering committee member of the standards organisation, the Coalition for Content Provenance and Authenticity (C2PA), a Joint Development Foundation project within the Linux Foundation. Adobe co-chairs the steering committee which meets regularly, chairs the Technical Working Group and has representatives on the User Experience Task Force, Threats and Harms Task Force. The C2PA also receives support from Adobe employees in Communications and Policy for C2PA external engagement.							

	We are committed to working with other C2PA members such as Microsoft, BBC, Intel, Google, Sony, Amazon, OpenAl and TruePic to ensure open technical standards for provenance are maintained to the highest standards; used to develop and implement content provenance across the ecosystem which is interoperable; and ultimately adopted by international standards organisations as the leading industry standard for digital content provenance.
	Internally, at Adobe we also have a team of full-time employees dedicated to working on provenance. This includes engineers helping to develop and maintain our open-source tooling for the community, user experience designers, and a team dedicated to advocacy and education, supporting adoption, and growing the community globally.
	In popular Adobe creative tools including Photoshop and Lightroom, consumers have access to attach Content Credentials to their digital content. Content Credentials are a free, open-source technology leveraging the C2PA open technical standard and serves as a "nutrition label" for digital content. Content Credentials can include important information such as the creator's name, the date the content was created, what tools were used to create an image and any edits that were made along the way.
	Other applications in Adobe Creative Cloud including Illustrator, Adobe Express, Adobe Stock, and Behance also support Content Credentials, and Adobe is continuing to roll out support for Content Credentials across products.
	Additionally, Adobe automatically attaches Content Credentials to content wholly generated with Adobe Firefly, the company's family of creative generative AI models, to ensure transparency around AI-generated content. This level of transparency allows customers to see content with context and helps build a more trustworthy digital ecosystem.
Measure 20.2	[insert wording if adapted] Enabling industry adoption of Content Credentials
QRE 20.2.1 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters]
	In January 2024, OpenAI announced details around its initiatives to protect the integrity of global elections, including support the C2PA standard and Content Credentials for images generated by DALL·E 3, which is now available and then support in its new video model, Sora.

In February 2024, Meta announced updates to its standard for labelling Al-generated images, ahead of global elections in 2024 – which includes its plan to build on the C2PA's industry standard solution for adding provenance to content to ensure that people have transparency around the content they see online.

Google joined the C2PA steering committee in February 2024 and is actively exploring how to incorporate Content Credentials into its own products and services in the future.

In March 2024, BBC News announced its implementation of Content Credentials, which are now embedded to its images and videos to verify content provenance and authenticity

In Spring 2024, Sony released firmware updates to its Alpha 1, Alpha 7S III, Alpha 7 IV cameras, adding support for Content Credentials.

In May 2024, TikTok joined the CAI and C2PA and began labelling AI-generated content uploaded to its platform with Content Credentials attached, with more integration to follow. This makes TikTok the first major social media platform to support Content Credentials.

Amazon joined the C2PA steering committee in September 2024 and began attaching Content Credentials to images created using the company's Titan Image Generator v1 and v2, the company's foundation AI models for enterprise.

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	[insert wording if adapted]
Measure 37.2	[insert wording if adapted]
Measure 37.3	[insert wording if adapted]

Measure 37.4	[insert wording if adapted]
Measure 37.5	[insert wording if adapted]
Measure 37.6	[insert wording if adapted]
QRE 37.6.1 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters]
	Adobe has participated in the 2024 Taskforce Plenary meetings. In addition, Adobe is an active member of the GenAl subgroup and has participated in the virtual meetings of the subgroup.

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	[insert wording if adapted]
	Adobe has an EU Code of Practice Tiger Team, which is an internal, cross-functional team that meets regularly to discuss the implementation of the Code commitments and our reporting requirements. This work is overseen by the General Counsel.
QRE 38.1.1 [insert wording if adapted]	Outline relevant actions [suggested character limit: 2000 characters]



	Reporting	on the service's response during an election
Threats observed or anticipated	d at time of reporting: [sugg	gested character limit 2000 characters].
Mitigations in place – or planne	ed - at time of reporting: [su	uggested character limit: 2000 characters].
the information below prov	ides an accurate and compl	elevant to their particular response to the threats and challenges they observed on their service(s). They ensure that lete report of their relevant actions. As operational responses to crisis/election situations can vary from service to dered 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 po	blicies	
Policy	Changes (such as newly introduced policies, edits, adaptation in scope or implementation)	Rationale
		Scrutiny of Ads Placements
Outline approaches pertinent t	o this chapter, highlighting	similarities/commonalities and differences with regular enforcement.
Specific Action applied (with reference to the Code's	Description of intervention	n

relevant Commitment and Measure)	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)	Description of intervention			
	Indication of impact (at beginning of action: expected 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 (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)

Description of intervention

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 2024 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 pledged 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 Al-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 committed 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 Al-generated audio, video, and image content that is generated by users with models in scope of this accord.

	Indication of impact (at beginning of action: expected impact) including relevant metrics when available		
Empowering the Research Community			
Outline approaches pertinent to	o 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 (at beginning of action: expected impact) including relevant metrics when available		
	Empowering the Fact-Checking Community		
Outline approaches pertinent to	o 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		