U.S. Senate AI Insight Forum: Elections & Democracy
Written Statement of
Andy Parsons
Senior Director of the Content Authenticity Initiative (CAI)
Adobe Inc.
November 8, 2023

Leader Schumer, Senators Rounds, Young, and Heinrich, and distinguished members of the Senate, thank you for the opportunity to participate in this important discussion today. The AI Insight Forums has an instrumental leadership role in shaping the future of artificial intelligence, advancing responsible innovation, and restoring public trust in the content we view online. It's an honor to spend time with you today, one year out from the 2024 elections, and share more on the importance of digital content transparency, standards, and tools to help people navigate an AI-powered world.

Since our founding in 1982, Adobe has pioneered cutting-edge technologies, bringing the world tools like PDF, Illustrator, and of course, Photoshop. We've been incorporating AI into our tools for over a decade to help creators realize their potential. With Adobe's expertise in AI and image science, we recognized that we could play a unique role in helping lead to a solution where we can use technology to help us prove what's true.

Addressing Deceptive Content & Dangers of Deepfakes

With the launch of Adobe's own foundation model, the text to image generating Adobe Firefly, the power and ability to generate content in seconds using just the click of a button is easier and more accessible than ever. Over 3 billion images have been created using Adobe's AI model in just eight months. Generative AI is transforming the way we work, create, and communicate. But in the hands of bad actors, this technology raises concerns about the ability to produce and spread false content at mass scale. It takes only seconds for millions of people to see and believe a lie. And we know from neuroscience that the age-old adage is true: Seeing is believing. That gives Generative AI images and video an extra power to deceive, and the same goes for Generative AI audio. In an election year, people's trust in digital content could have substantial consequences.

AI based deepfakes and their consequences are beginning to show up everywhere. A video of Ukrainian President Zelensky ordering troops to surrender to Russian forces. A picture of an explosion at the Pentagon – which moved markets. These examples illustrate not only the ability of altered digital content to deceive us, it also shows how the potential of deepfakes can make us doubt what's real – threating civil discourse and democracy. Once people realize there is misinformation out there, they are at risk of not just believing lies but of no longer believing the truth. We need a way to distinguish fact from fiction, to give people the tools they need to know what they can trust.

While misinformation has taken on a new prominence in our political dialogue today, the problem is not new. Even in our nation's earliest elections, bad actors used misinformation as a tool to deceive voters and manipulate opinions. And as society has continued to develop new technologies, bad actors have tried to leverage these technologies to do harm.

Today, we are here to discuss one of the most critical areas when it comes to responsible AI development as we look to uphold democracy and protect our election process – we must address the dangers of deepfakes and align on the solutions needed to combat them.

How We Can Restore Trust: Digital Content Provenance Standards

We believe one of the most effective components of restoring trust in our digital world is to establish global standards that leverage the power of provenance to help creators prove what's true. And so, along with others, that is why we established an independent non-profit standards organization called the Coalition for Content Provenance and Authenticity (C2PA) to guide the development and implementation of provenance technologies.

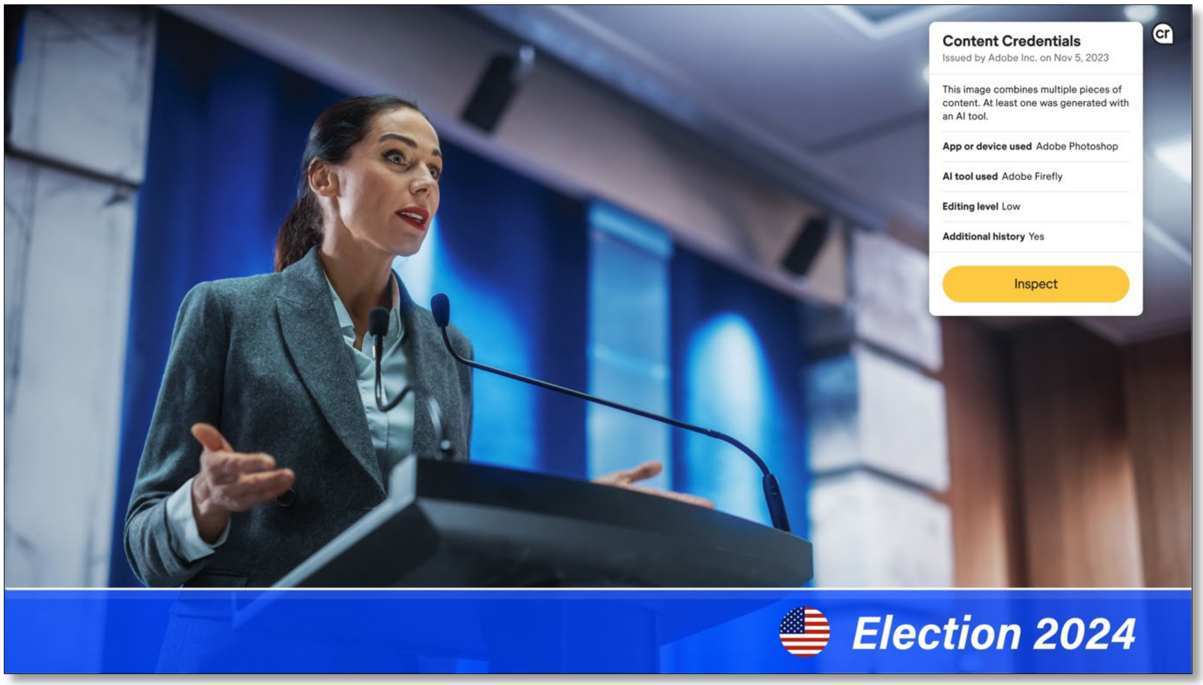
The C2PA publishes and promotes an open technical standard for provenance technology so that anyone can implement it into their tools and platforms. The standard started with images, but now incorporates solutions for audio, video and documents as well. The technology, called Content Credentials are built on this open technical standard, and the code for Content Credentials is free and available in an open-source toolkit. These tools enable the broader developer community to integrate content provenance across web, desktop and mobile project implementations for creators and consumers alike. Anyone can build Content Credentials; everyone should.

Previous efforts to combat deepfakes have focused on automated detection of deepfakes.
However, we believe detection is not the answer. Detection technologies face significant limitations: the high error rates in incorrectly classifying fake media are too high to be relied upon by the public. And the challenge with deepfake detection is that it often happens after the fact – and too late. By the time you attach a label to a lie, millions of people have already seen it and assumed it's credible. You can't retract lies at scale.

Content Credentials: How They Work

Leveraging the power of provenance, Content Credentials function like a nutrition label for digital content: they allow creators to attach information to their content such as name, date, and what tools were used, as well as edits that were made along the way. The information travels with a piece of content wherever it goes. A simple icon accompanies the content and lets viewers know that there is more information about that content for them to see.

Once clicked, a panel shows up with just the critical metadata a viewer may be interested in, such as the creator and whether generative AI was used. If someone ones to see rich, contextual detail, an entire website has been created to allow a user to drag an image to see its Content Credentials and easily compare the content with its original to see what changes were made. This way, Content Credentials give people a way to show their work and give consumers a way to see context alongside the content they are consuming.



Content Credentials

Issued by Adobe Inc. on Nov 5, 2023

This image combines multiple pieces of content. At least one was generated with an AI tool.

App or device used Adobe Photoshop

AI tool used Adobe Firefly

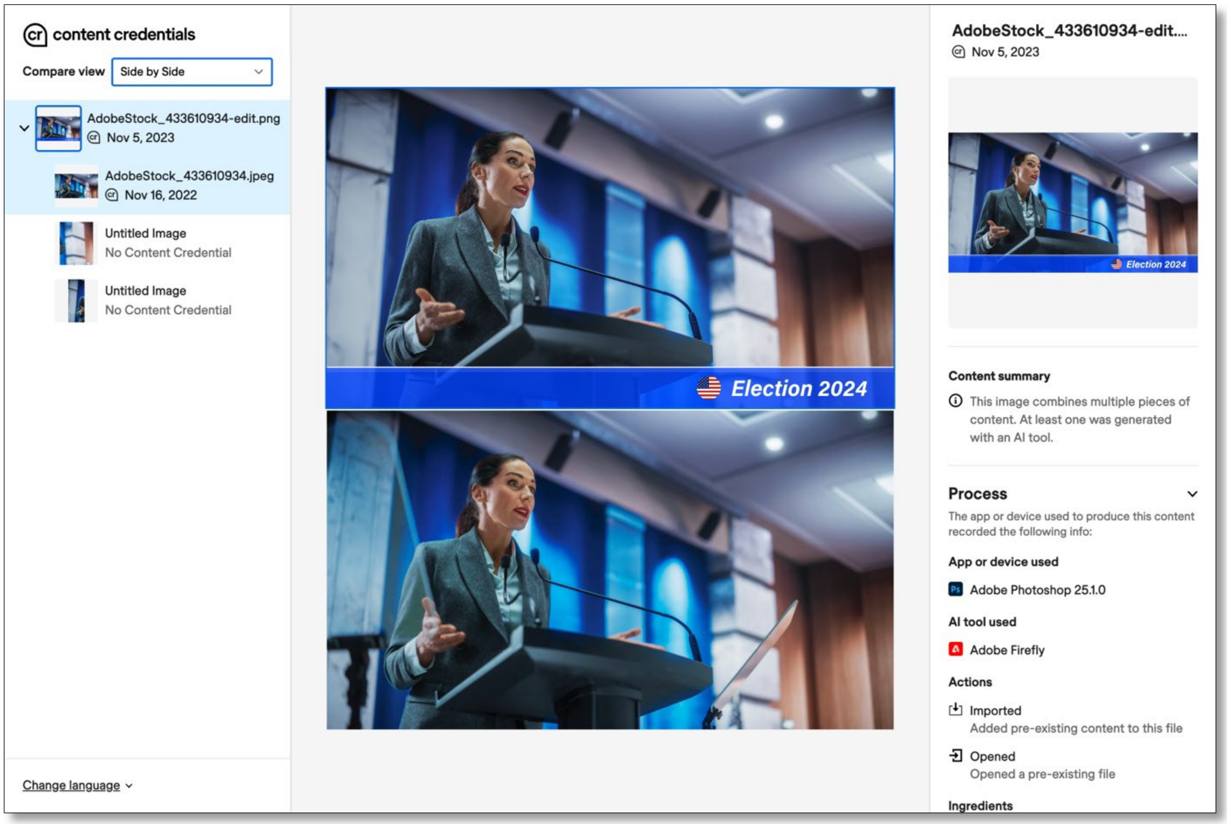
Editing level Low

Additional history Yes

Providing context is critical. Soon, the majority of content will have some AI generated pixels in it and almost all will have used AI assisted editing tools. Simply labeling content as “AI” won’t help the user understand if they should trust the content or not. Let’s say a candidate wants to use content of herself shaking hands with voters. The candidate may want to use AI to remove distracting background elements to focus the viewer on the important aspect of the photo or video. Even though content was edited with AI, the image itself is still authentic for the purposes for which it is being shown – the candidate was there and was shaking hands with voters. This type of journalistic editing happens everyday. This is why we need to label the content with provenance, to let people know who created it, signed it, and give them a way to easily see the changes that were made, so they can decide for themselves whether to trust it.

Content Credentials are opt-in and are designed to be tamper evident. Creators can choose what information they want to share, and, for privacy reasons, whether for the fear of reprisal or a desire to be anonymous, some may choose not to attach a credential or limit what information is shown. The level of transparency provided will impact the level of trust readers will place in the content, but that choice will be up to the creator. The decision to trust is in the hands of the public; with the Content Credentials approach, tech companies and governments will not be the arbiters of truth.

Content Credentials are based on digitally signed information, indelibly part of the content, and if any tampering occurs along the way, the tampering will be evident so the public will know to be skeptical. To ensure the robustness of this solution, the standard supports imperceptible watermarking, placed in the media itself, which can link back to the original Content Credential if it was stripped out. Finally, Adobe created a free public website called Verify at https://contentcredentials.org/verify. Here, you can bring an image, video, or audio file with Content Credentials to see its full history in one easy interface. If someone tried to strip out the Content Credentials, for example by screenshotting or photographing a screen, the Verify site has matching technology that can recover them if the Content Credentials were stored in the cloud. That way, you can always go back to the source to see what really happened.



Content summary

This image combines multiple pieces of content. At least one was generated with an AI tool.

Process

The app or device used to produce this content recorded the following info:

App or device used: Adobe Photoshop 25.1.0

AI tool used: Adobe Firefly

Actions

Imported
Added pre-existing content to this file

Opened
Opened a pre-existing file

Ingredients

In order for this solution to truly work, we need it everywhere we create and consume digital content. That is why we co-founded the Content Authenticity Initiative (CAI), to promote adoption of the C2PA open standard and content credentials. The CAI now counts more than 2,000 members, including chipmakers like NVIDIA and Qualcomm, software and AI developers such as Adobe and Microsoft, camera manufacturers like Leica, Nikon and Canon, news organizations like the Associated Press, Reuters, National Geographic, the Wall Street Journal, BBC, and more. Recently, Universal Music Group joined the CAI because they recognize the importance of this type of solution across all types of media, including images, video, and audio.

The recent White House Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence is an important step forward in the journey to create a responsible AI framework. Requiring federal agencies to use provenance to label the content they create will help establish the authenticity of the information they are sharing with their citizens and will also provide leadership for the need to get authenticity adopted as a standard globally.

It is imperative to have a provenance solution everywhere. Culturally, we need to educate people that first, they can't trust everything they see and hear digitally. Then, we need to inform them that there are tools they can use to verify for themselves whether or not to trust the content. Then, people will expect that important news stories or factual events will come with provenance before they believe it. And when people see important content without it, they will know to be skeptical. That is the state we must achieve in order to restore trust in digital content.

With the 2024 elections drawing nearer, we are at a pivotal moment. With the right tools and frameworks, like Content Credentials and the C2PA open-source standard, we can create a trustworthy digital space for everyone.

Recommendations:

- **Require provenance for all digital campaign content:** We believe Content Credentials technology and the C2PA open standard is part of the solution to help create trust in election content. To that end, Congress should consider legislation requiring campaigns to integrate Content Credentials and digital provenance into their online campaign communications. This would afford voters a level of transparency about what they are consuming and be better equipped to make an informed decision to trust the material or not. If they see the content with Content Credentials, they can decide to trust it. If the content does not have it, they will know to be skeptical. And anyone creating content for the purposes of deceiving the public can be held liable for the misuse of digital editing technology. We also believe that because the Federal Election Commission (FEC) has the authority to regulate against fraudulent misrepresentations in campaigns, that they have an important role to play too in addressing the problem deep fakes in our elections process. In fact, we recently recommended to them in their petition for rulemaking concerning AI in campaign ads that they require Content Credentials, or similar provenance technology, in all federal campaign materials and advertising to provide greater transparency and trust to voters.
- **Require provenance metadata to be maintained wherever it goes:** Today, even if a campaign creates content with Content Credentials, a platform might strip the metadata away. Congress should require carrying provenance wherever it goes to ensure the public can see it wherever they are consuming online content. We additionally provided this recommendation to the FEC.
- **Support for state and local governments: Publish national standards and best practices:** Our elections are not administrated by a unified national system. Rather, our elections are managed by dedicated public servants across thousands of state and local systems. However, the Federal government can play an important role in publishing resources, best practices and promulgating guidelines for state and local governments to follow including encouraging them to use and require provenance in election campaign content. Federal agencies like the Cybersecurity Infrastructure Security Agency (CISA) and the Election Assistance Commission (EAC), or non-profits like the Center for Internet Security (CIS) serve an important role in this regard, publishing best practices and other guidance about how states and local governments can secure their elections.
- **Set an example through adoption:** We believe Congress should set the standard for applying this level of digital content transparency in their own communications. The House and Senate should require its own digital content to have provenance embedded, to help establish trust with the public in the content they are consuming.

Conclusion: Safeguarding elections in the age of AI: The time to act is now

With the 2024 elections drawing nearer, it is critical we align on a content provenance standard and approach and dedicate specific focus to election content, given how high the stakes are. We are at a pivotal moment. AI has the power to unleash human creativity in new ways, raise human ingenuity to new levels, and lift our society to unimagined heights. But it has to be done right. With the right tools and frameworks, we can create a trustworthy digital space for everyone.

###