Making Deepfakes Gets Cheaper and Easier Thanks to A.I.

The New York Times

The Battle Against Fake News Enters The Age Of Deepfakes

Forbes

Puffer coat Pope. Musk on a date with GM CEO. Fake AI ‘news’ images are fooling social media users

CNN Business

Elections in UK and US at risk from AI-driven disinformation, say experts

The Guardian

From Scams to Music, AI Voice Cloning Is on the Rise

TIME

With rise of AI-generated images, distinguishing real from fake is about to get a lot harder

CBC News
Membership ~2500

CBC Radio-Canada  EFE  dpa  GANNETT  gettyimages
infobae  shutterstock  REUTERS  epaimages  stern  EL TIEMPO
BBC  axel springer  AP  AFP  NATIONAL GEOGRAPHIC SOCIETY  npr
Dentsu  PUBLICIS GROUPE  OmnicomGroup  truepic  arm  camera bits
Canon  Microsoft  Qualcomm  Nikon  Leica  Adobe
Pillars of Transparency and Authenticity

Provenance

Education

Detection

Policy
Provenance

The basic trustworthy facts about the origins of a piece of digital content (image, video, audio recording, document)
Provenance

Proving what is real (as opposed to detecting what is false)

- Creators
- News media
- AI-generated content
- Brand reputation
- E-Commerce
- Insurance
- Auditing
- Law enforcement
- Medical/Scientific
- Satellite imagery
- Election Integrity
Content Credentials

1. Capture
Work with manufacturers to integrate CAI into smartphones and cameras.

2. Edit
Integrate CAI into editing tools, both Adobe products and others.

3. Publish
Publishing systems maintain CAI metadata throughout their platforms.

4. Trust
Clear & universal user experience reveals provenance information.
Content Credentials

1. **Capture**
   Work with manufacturers to integrate CAI into smartphones and cameras.

2. **Edit**
   Integrate CAI into editing tools, both Adobe products and others.

3. **Publish**
   Publishing systems maintain CAI metadata throughout their platforms.

4. **Trust**
   Clear & universal user experience reveals provenance information.

Viewing content credentials at any point in the content’s journey shows what (if anything) was done to change the asset.
Content Credentials

- **Creation**
  - Creation date, Created with, Creator

- **Editing**
  - Edit date, Edits made, Edits with, Editor

- **Publishing**
  - Review, Publish date, Published by

- **Sharing**
  - Optimized for delivery

- **Viewing**
  - Provenance
content credentials
What about AI?
Wait, where did this image come from?

Deepfakes. Voice cloning. Synthetic media. It’s hard to tell what’s accurate and authentic these days.
Invalid video Content Credentials or assets

This sample shows how the Content Credentials player can surface important information to viewers when there are issues with a video's Content Credentials or segments within a video.
## Secure Metadata

<table>
<thead>
<tr>
<th>Assertions</th>
<th>Asset Hashes</th>
<th>Content Credential</th>
</tr>
</thead>
<tbody>
<tr>
<td>What was done to an image, by whom, when etc.</td>
<td>Cryptographically bind assertions to image using unique codes</td>
<td>Bundle of assertions, asset hashes &amp; digital signatures</td>
</tr>
<tr>
<td>File</td>
<td>Cloud</td>
<td>Blockchain</td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>------------</td>
</tr>
<tr>
<td>Can be stored in the file</td>
<td>Stored in the cloud - with pointer from file to cloud. Thumbnail of file stored for recovery of Content Credential</td>
<td>Distributed ledger</td>
</tr>
</tbody>
</table>
The Coalition for Content Provenance and Authenticity (C2PA)

- Standards organization
- Within Linux Foundation
- Adobe on Steering Committee & Chairs TWG
- V1.0 C2PA Specification Released on Jan 2022
All have announced implementation
Open-source tool 1: JavaScript UI Kit

Implementors will use to
Display Content Credentials on your site or app
Link Content Credentials displayed on your site or app to Verify

Install to create, verify and explore content credentials on your command line. Or wrap it into a service to quickly equip your processes to interact with content provenance.
Open-source tool 2: C2PA Tool

Implementors will use to

Display Content Credentials on your site or app

Link Content Credentials displayed on your site or app to Verify

Write Content Credentials data into files

Quickly create and inspect Content Credentials data

Everything you need to develop rich, browser-based experiences with content credentials.
Open-source tool 3: Full SDK

Implementors will use to

Display Content Credentials on your site or app

Link Content Credentials displayed on your site or app to Verify

Write Content Credentials data into files

Quickly create and inspect Content Credentials data

Customize displaying and creating Content Credentials data, with the full power of the specification

Deploy on Web, mobile, and desktop

Develop custom applications across desktop, mobile, and services that create, verify, and display content credentials via our powerful Rust library.
3 key areas of focus

1. Open Standards Specifications
2. Implementation & Member Collaboration
3. Advocacy & Education