Overview

Content Authenticity Initiative

Kate Brightwell

Head of Policy and Government Relations UK & Europe, Adobe

Why We’re Here

- Responsible Use
- Intent to Deceive
- Lack of Transparency, Trust, Attribution
- Generative tools:
  - DALL-E
  - Midjourney
  - Stable Diffusion (open source)

A deepfake video of Mark Zuckerberg presents a new challenge for Facebook

By Rachel Metz and Donie O'Sullivan, CNN Business

Updated 10:11 PM ET, Tue June 11, 2019

A Military Coup in Gabon Inspired by a Potential Deepfake Video is Our Political Future

Ty Joplin

Published May 8th, 2019 - 09:54 GMT

'Deepfake geography' could be the latest form of online disinformation

COMMENTS

By Hebe Campbell & Matthew Holroyd • Updated: 07/05/2021 - 19:48

Hong Kong isn't just battling on the streets: There is also a war on misinformation online

By Jessie Yeung, CNN

Updated 7:47 PM ET, Sun August 11, 2019

The viral Mars Perseverance rover video going around is fake

Mark Kaufman · 2/20/2021

CBS NEWS / *May 25, 2019, 12:39 PM*

Doctored Nancy Pelosi video highlights threat of "deepfake" tech

The coming deepfakes threat to businesses

Kaveh Waddell, Jennifer A. Kingson Jul 19, 2019

Beijing's new weapon to muffle Hong Kong protests: fake news

India WhatsApp rumors: Mob kills man in latest attack, 30 arrested

By Sugam Pokharel and James Griffiths, CNN

Updated 5:09 AM ET, Mon July 16, 2018

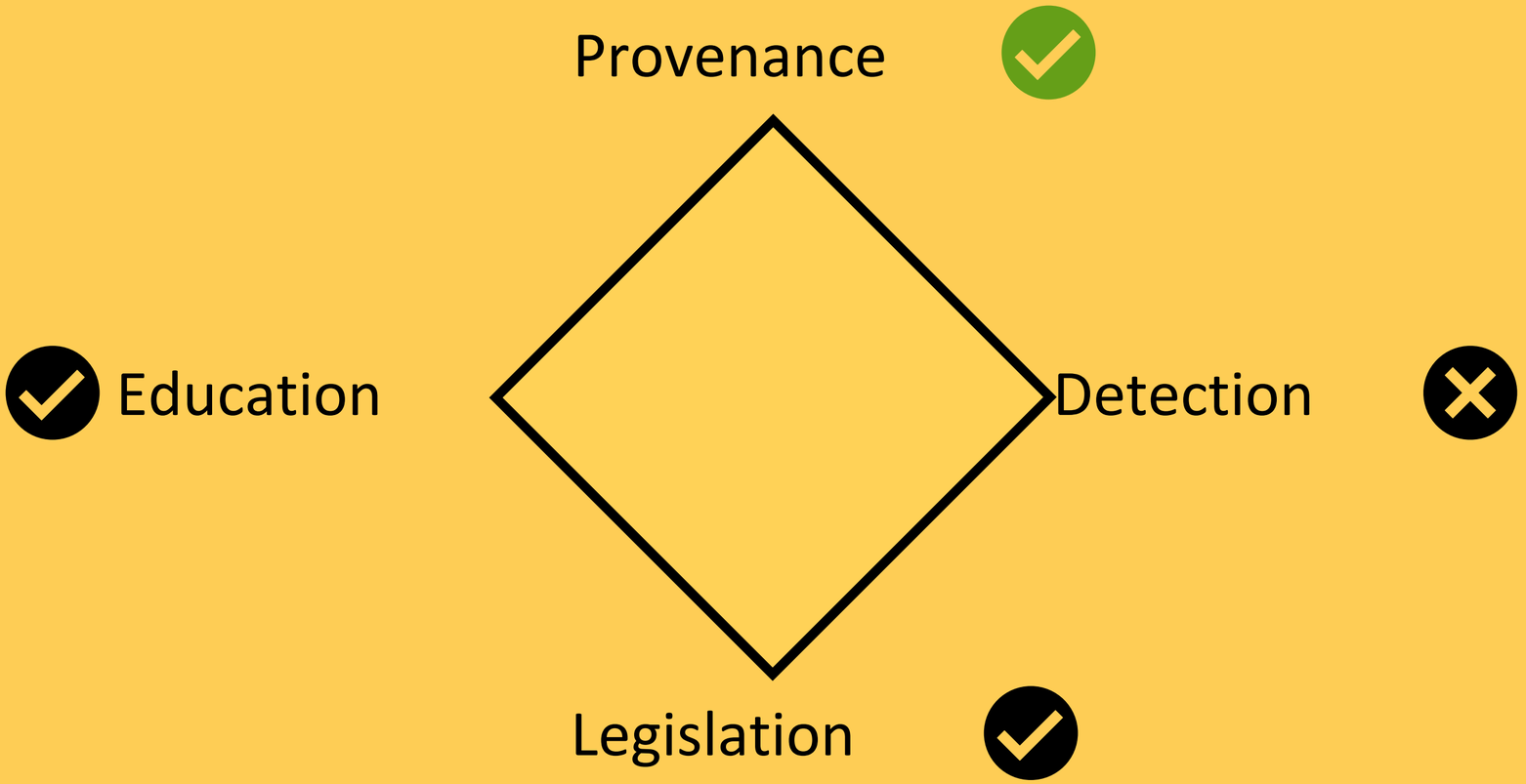
Why Not Detection?

Detection is likely to fail because it’s an arms race.

The 2020 Deepfake Detection Challenge achieved only a 65% hit rate.

Generative engines like DALL-E 2 and Midjourney can produce photorealistic images from textual input.

The fight against mis/disinformation



What is Provenance?
Where We Are Now
The Road to Adoption

Content
Authenticity
Initiative

What is Provenance?

Content
Authenticity
Initiative

Digital Provenance is the ground truth about how an asset was: Made, Altered, Delivered.

Instead of guessing what's false, we can prove what's true.

Content Authenticity Initiative

Creation

= Creation date, Created with, Creator

Editing

= Creation date, Created with, Creator + Edit date, Edits made, Edits with, Editor

Publishing

= Edit date, Edits made, Edits with, Editor + Review, Publish date, Published by

Sharing

= Review, Publish date, Published by + Optimized for delivery

Viewing

= Creation date, Created with, Creator + Edit date, Edits made, Edits with, Editor + Review, Publish date, Published by + Optimized for delivery

Key Product Concepts

- Make it *simple* for creators and organizations to capture and publish provenance.
- Make it *intuitive* and *expected* for content consumers to view provenance.

Guiding Principles

- Privacy
- Global Accessibility
- Protections Against Misuse
- Fit with Existing Workflows
- Simplicity and Minimal Cost
- Interoperability
- Extensibility

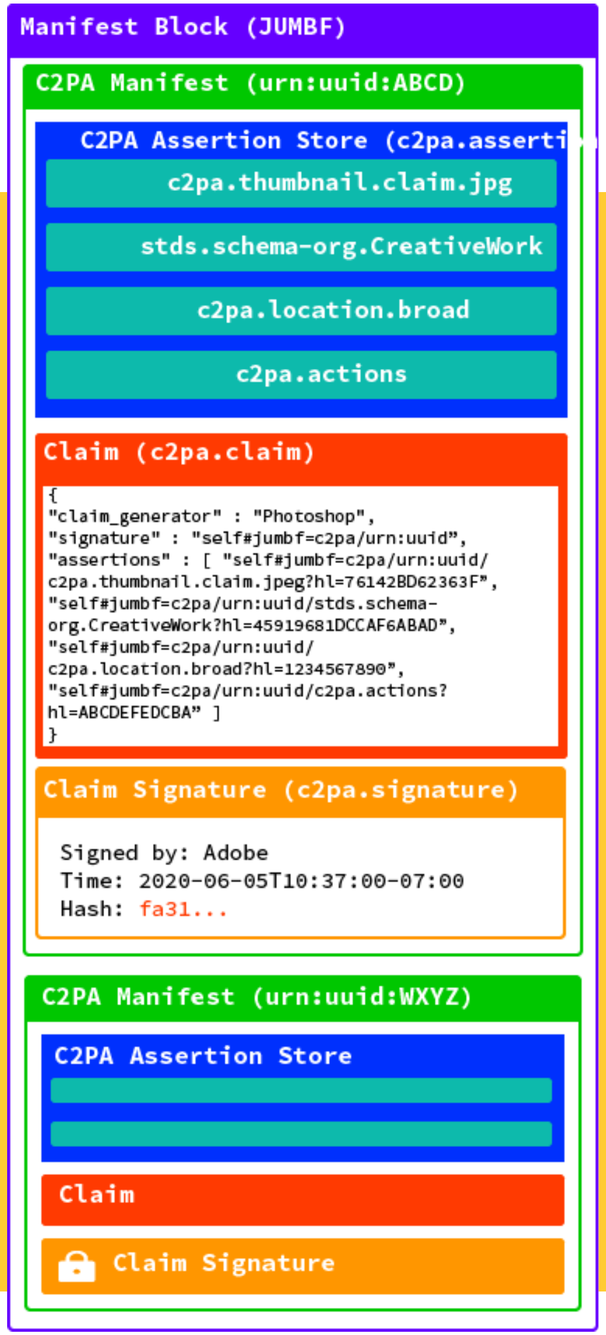
Provenance in Practice

Provenance-enabled media has a *Manifest*:

- Contains captured assertions about the media
- Can be associated with the creator's identity or organization
- Is cryptographically bound to the media and signed

If you have media with a *Manifest*, you can be certain:

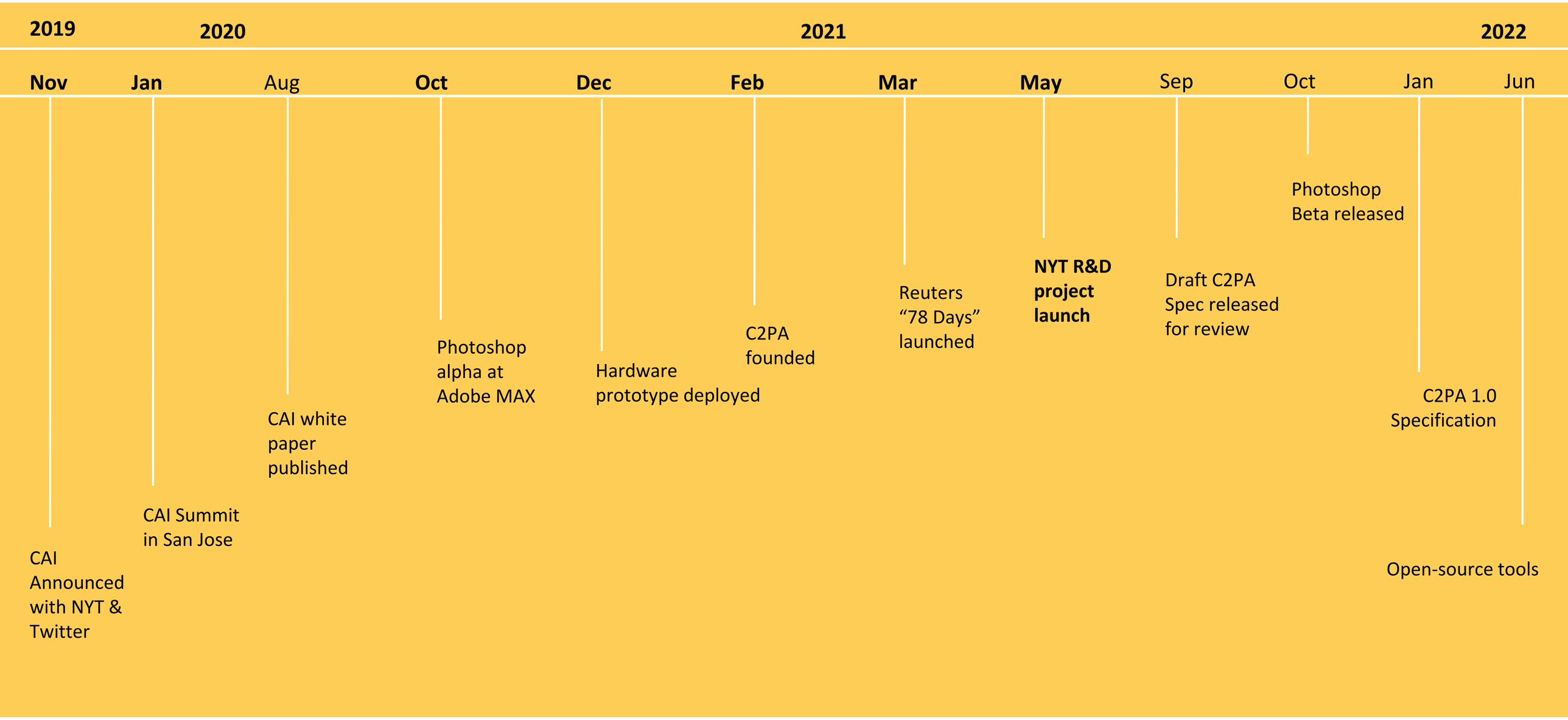
- It came from the person or organization that signed it
- It has not been altered

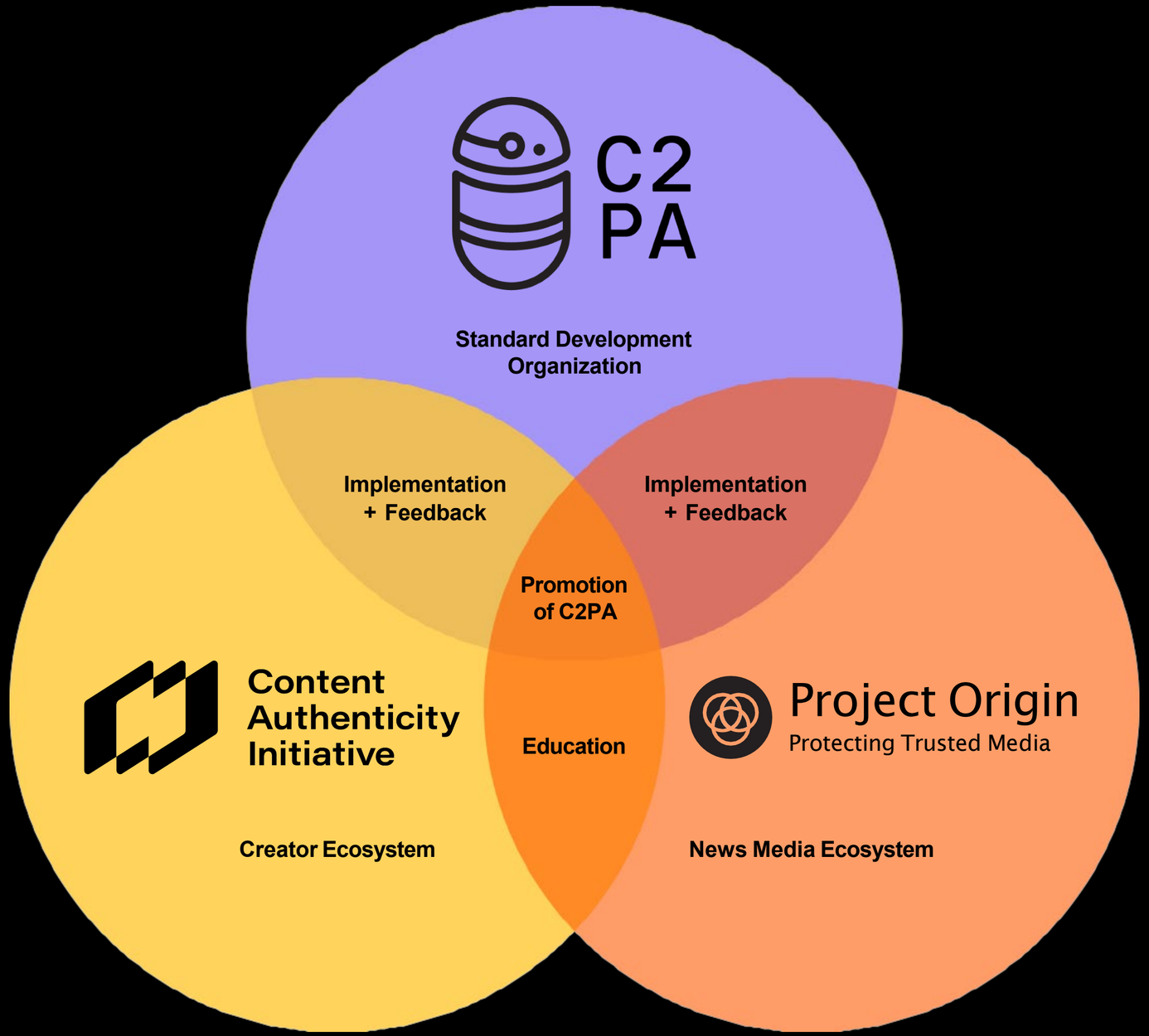


Where we are now

Content
Authenticity
Initiative

Timeline


2019
2020
2021
2022
Nov
Jan
Aug
Oct
Dec
Feb
Mar
May
Sep
Oct
Jan
Jun
CAI Announced with NYT & Twitter
CAI Summit in San Jose
CAI white paper published
Photoshop alpha at Adobe MAX
Hardware prototype deployed
C2PA founded
Reuters “78 Days” launched
NYT R&D project launch
Draft C2PA Spec released for review
Photoshop Beta released
C2PA 1.0 Specification
Open-source tools


C2PA
Standard Development Organization
Implementation + Feedback
Implementation + Feedback
Promotion of C2PA
Education
Content Authenticity Initiative
Creator Ecosystem
Project Origin
Protecting Trusted Media
News Media Ecosystem

Coalition for Content Provenance and Authenticity (C2PA)

Adobe

Microsoft

arm

BBC

Truepic

intel

SONY

- Run under the Linux Foundation
- Spec developed in < 1 year
- v1.0 released in January 2022
- Images, audio, video (to start)
- Threats/Harms + UX Conventions

Content Authenticity Initiative

Membership 800+

Content Authenticity Initiative

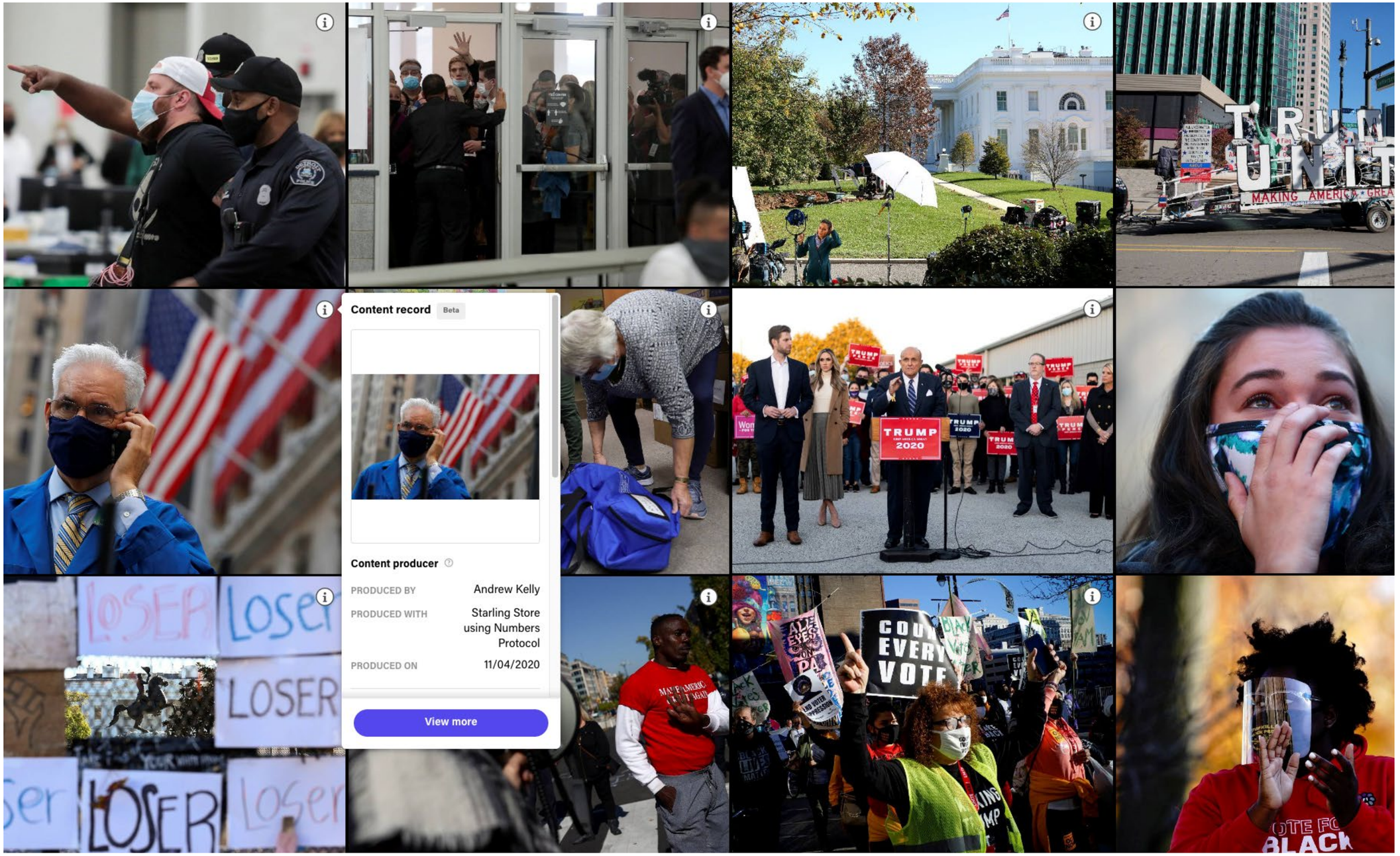
4 key areas of focus:

- Standards specifications via the C2PA
- Deployment through member collaboration and Adobe products
- Education and advocacy
- Open-source tools

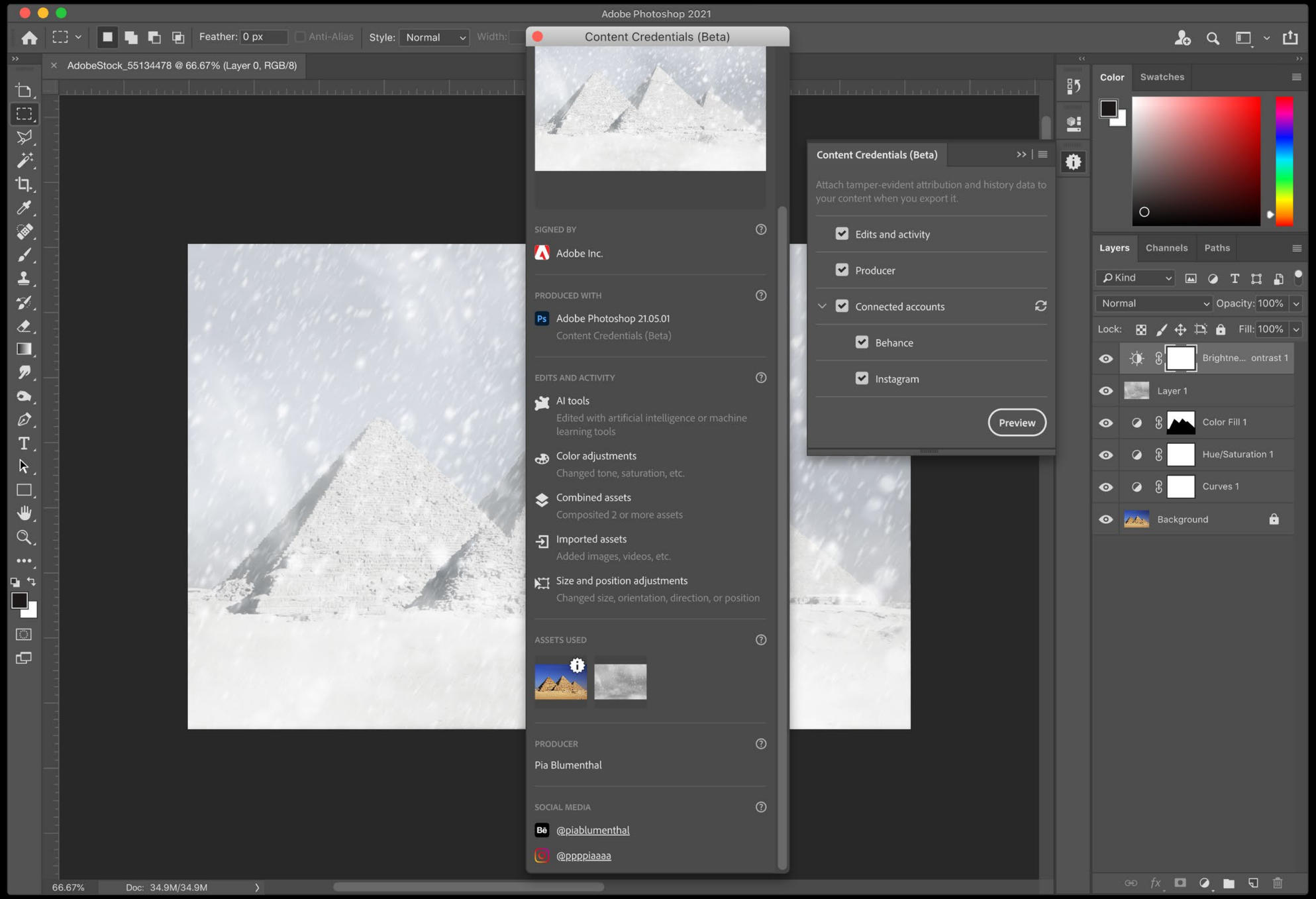
Examples

Content
Authenticity
Initiative

Reuters &
Starling Lab
“78 Days”



Content Credentials in Photoshop



The Road to Adoption

Content Authenticity Initiative

Broad adoption of the C2PA standard requires participation.

- Camera hardware
- Mobile phone hardware
- Social media platforms
- Publishing systems
- Operating systems

Participants in the C2PA, CAI, and Project Origin communities are committed to deploying the standard in products. The ecosystem is growing day-by-day.

Role of Regulators

- Media literacy duties – provenance can empower users
- Consideration of platform and publisher adoption and risk assessments/incentives
- Advisory boards looking beyond detection and including industry experts
- Ability to recommend open technical standards

Content
Authenticity
Initiative

Thank you.

kbrightwell@adobe.com

contentauthenticity.org

opensource.contentauthenticity.org

c2pa.org

@contentauth