

Examining the Impact of Provenance-Enabled Media on Trust and Accuracy Perceptions



K. J. Kevin Feng
University of Washington



Nick Ritchie
BBC UX&D



Pia Blumenthal
Adobe



Andy Parsons
Adobe



Amy X. Zhang
University of Washington

CSCW 2023
October 14-18, 2023 | Minneapolis, MN



Social
Futures
Lab

DUB
DESIGN
USE
BUILD

W C2
PA



juju 💰 ✅
@ayeejuju

Subscribe

...

only florida would have sharks swimming around on the freeway during a hurricane 🤦‍♂️🤦‍♂️



4:20 PM · Sep 28, 2022

73

761

18.8K

172



Post your reply

Reply



juju 💰✓
@ayeejuju

Subscribe

...

only florida would have sharks swimming around on the freeway during a hurricane 🤣🤣



4:20 PM · Sep 28, 2022

73

761

18.8K

172

↑



Post your reply

Reply



Jules
@crownjulesss

Believe it or not, this is a shark on the freeway in Tseung Kwan O, Hong Kong. #SuperTyphoonMangkhut #Mangkhut2018



10:47 PM · Sep 15, 2018

6

41

75

↑



Jym
@JymFox

There's a SHARK on the FREEWAY in PHILADELPHIA RIGHT NOW



8:38 PM · Feb 4, 2018

3

6

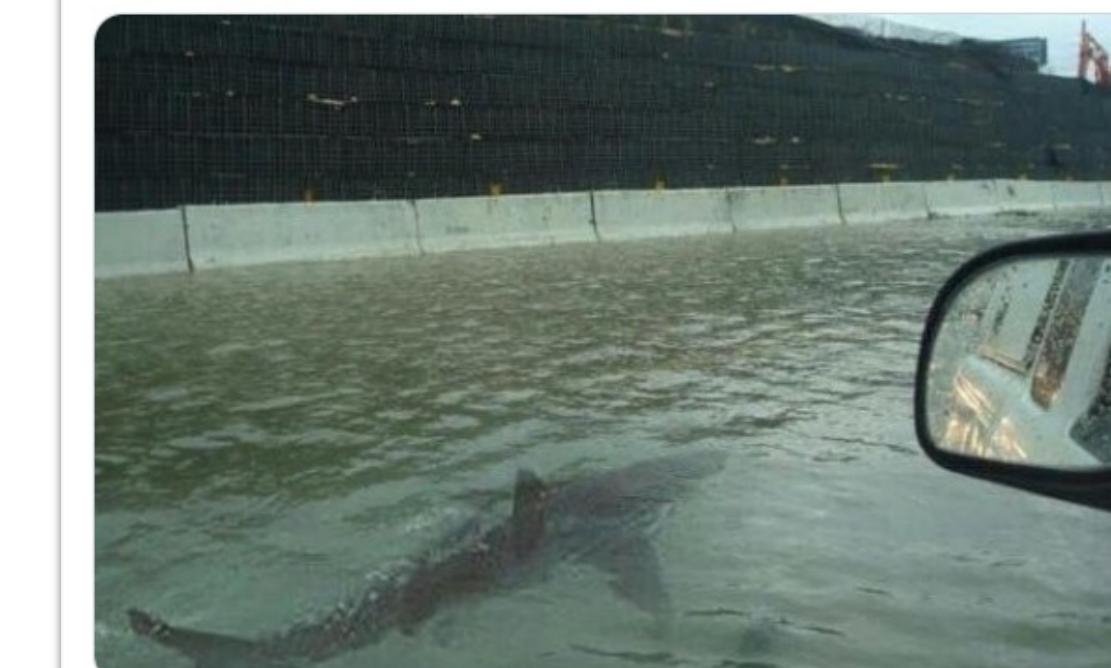
25

↑



Trolly McTrollsalot
@McTrollsalot

Believe it or not, this is a shark on the freeway in Mexico Beach, Florida. #HurricaneMichael #MichaelHurricane2018 #HurricaneMichael2018 #SharkHamill



12:21 PM · Oct 10, 2018

13

19

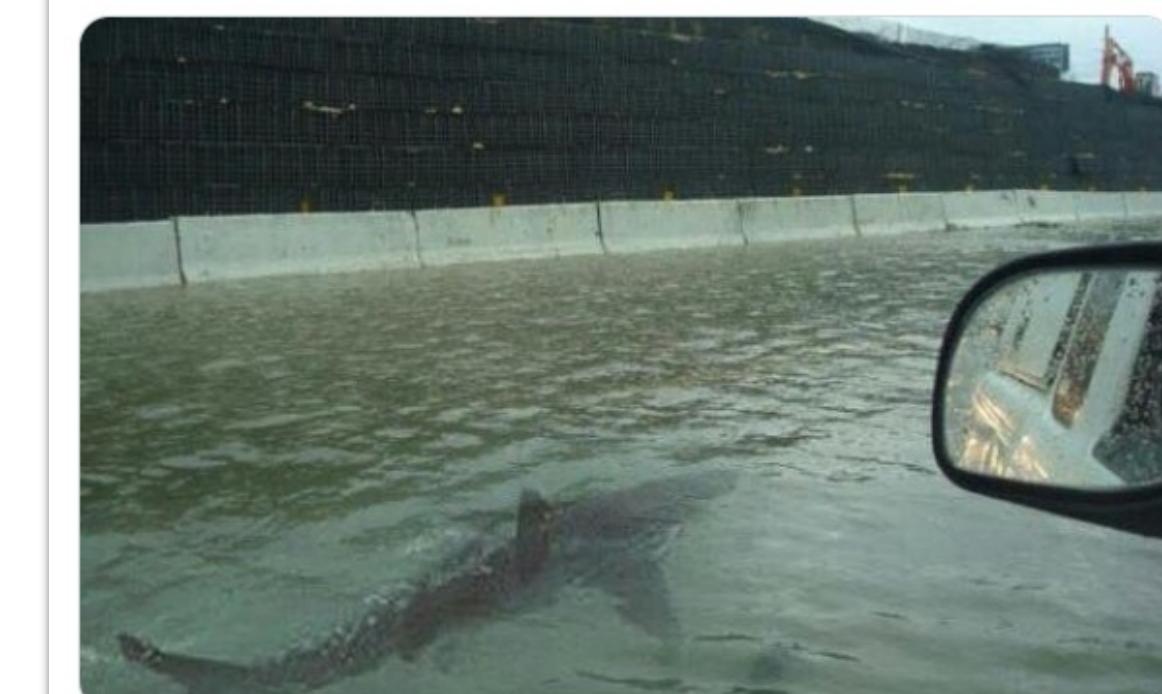
21

↑



calebquinoa39 7
@FortyUter

Believe it or not, this is a shark on the freeway in Salt Lake City, Utah. #HurricaneHary



3:19 PM · Aug 28, 2017

8

4

47

↑

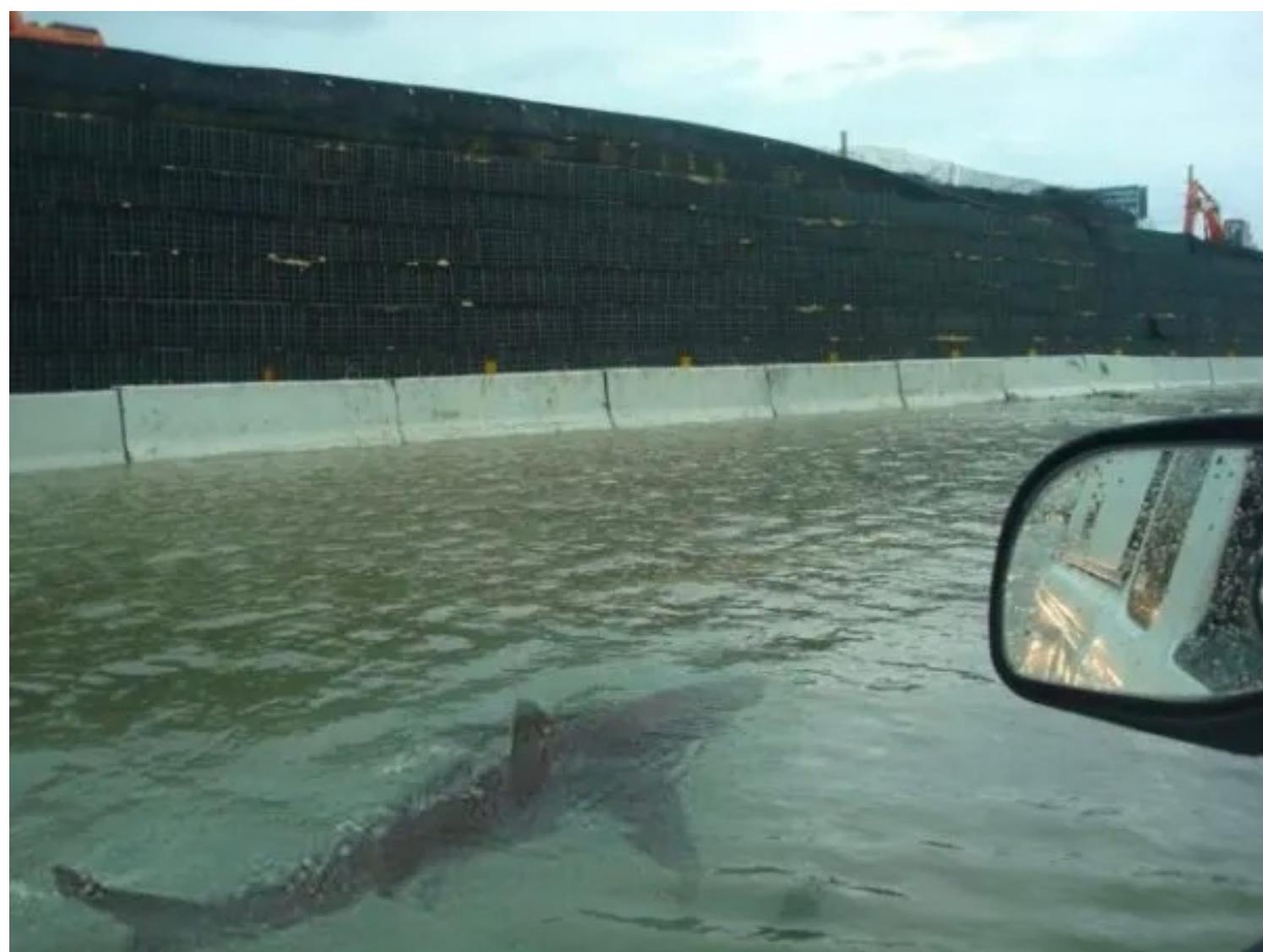
[Credit: Snopes Fauxtography archives]



Africa Geographic, September 2005



Reddit & Twitter, ~2011



 Twitter

Sep 8, 2021 at 10:34 AM



 Adobe

Sep 6, 2021 at 8:22 AM

PRODUCED BY

John Smith

EDITS AND ACTIVITY

 Color adjustments

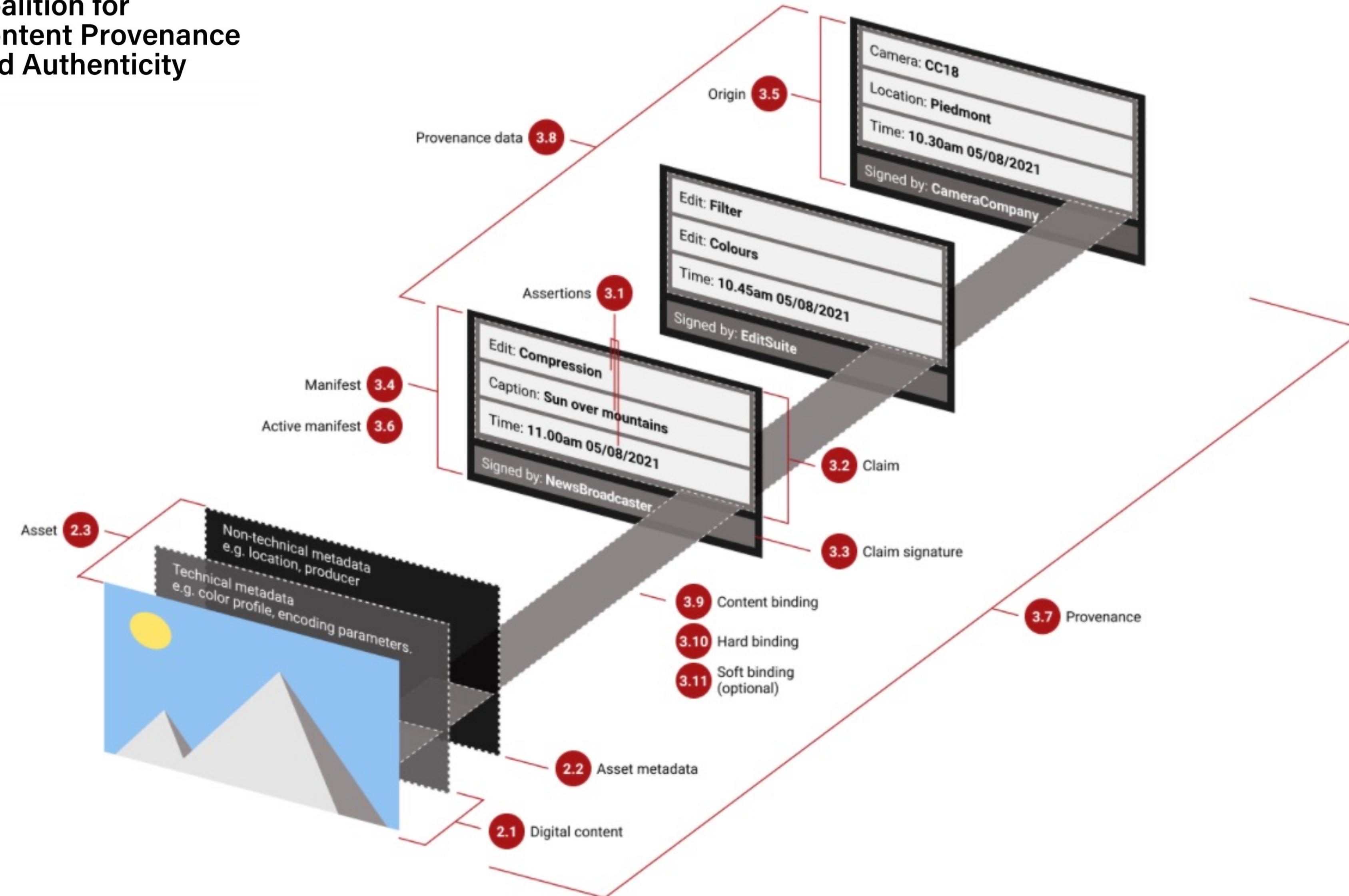
 Combined assets



[View more](#)



Coalition for Content Provenance and Authenticity



C2PA Technical Standard (<https://c2pa.org>)



Introduction

A Provenance Chain

Media journey

Publish



Media was published by Twitter

Edit



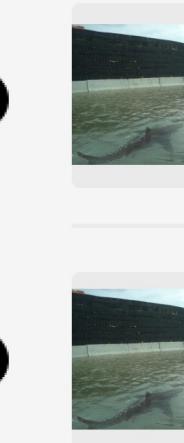
Media was composited in Adobe Photoshop

Capture/ Create



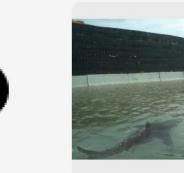
Multiple origins

Provenance UI



Twitter

Sep 8, 2021 at 10:34 AM



Adobe

Sep 6, 2021 at 8:22 AM

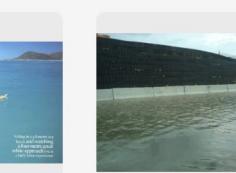
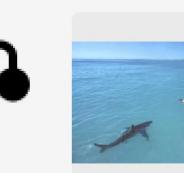
PRODUCED BY

John Smith

EDITS AND ACTIVITY

Color adjustments

Combined assets



[View more](#)

Introduction

A Provenance Chain

Media journey

Publish



Media was published by Twitter

Edit



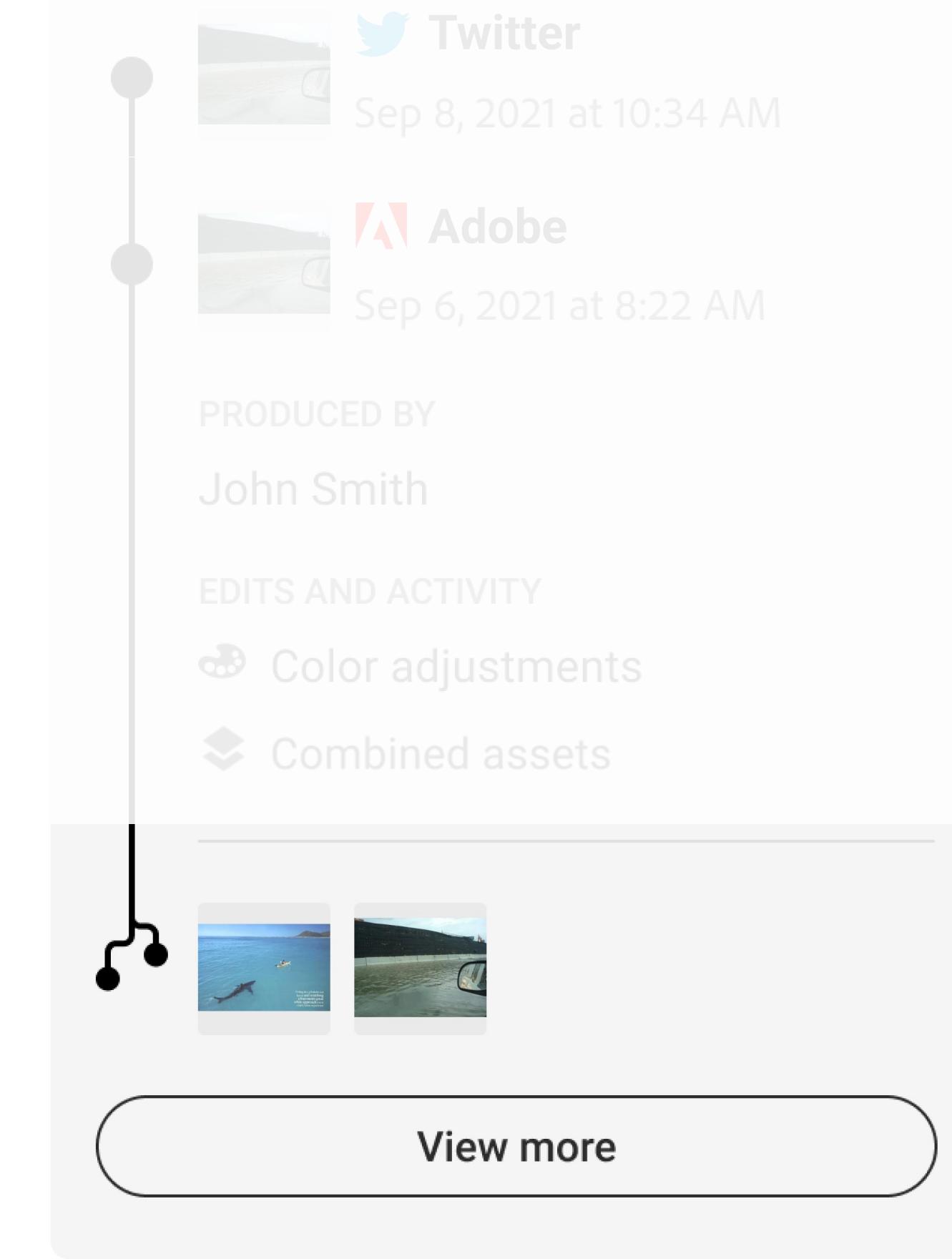
Media was composed in Adobe Photoshop

**Capture/
Create**



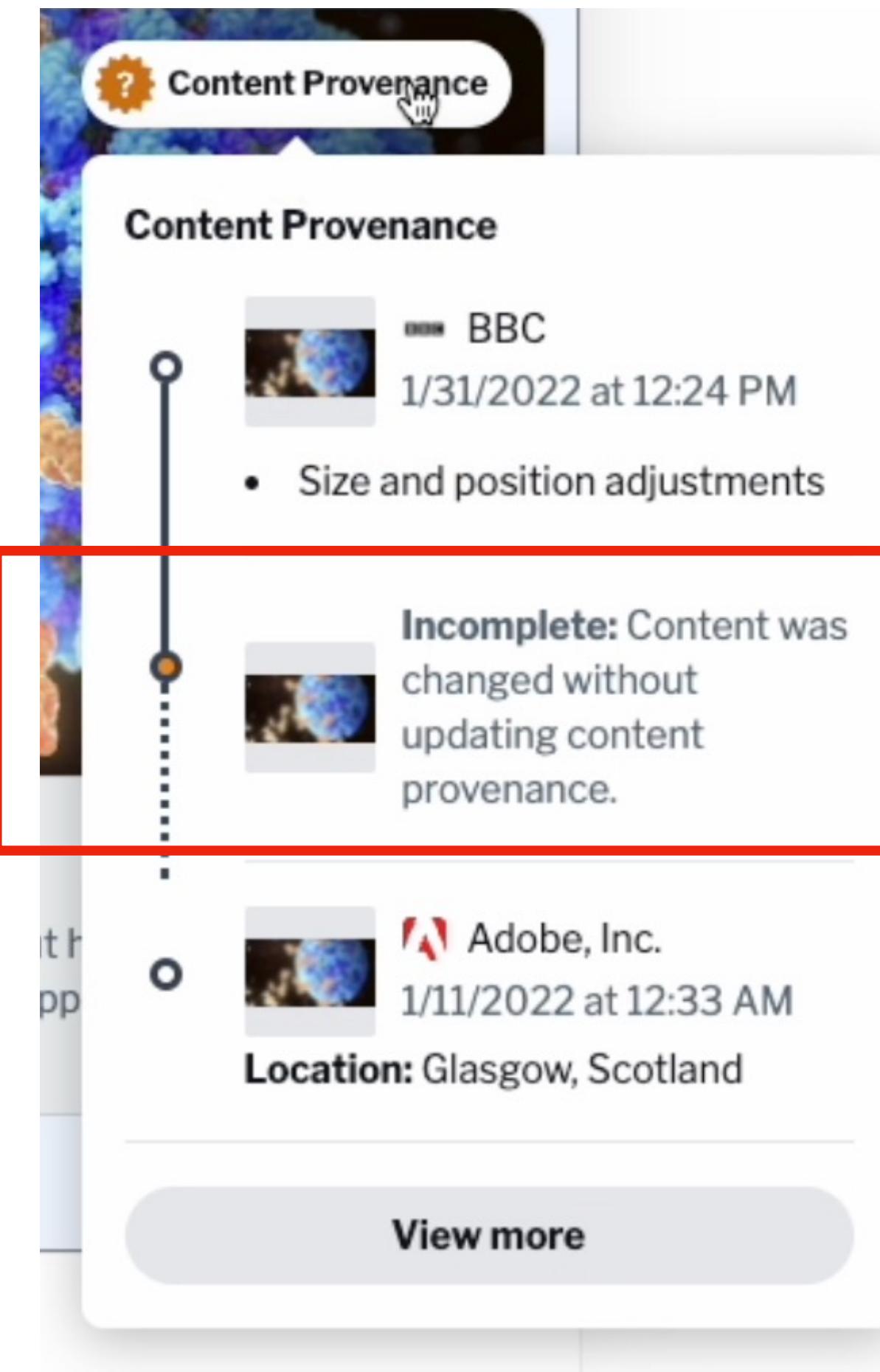
Multiple origins

Provenance UI

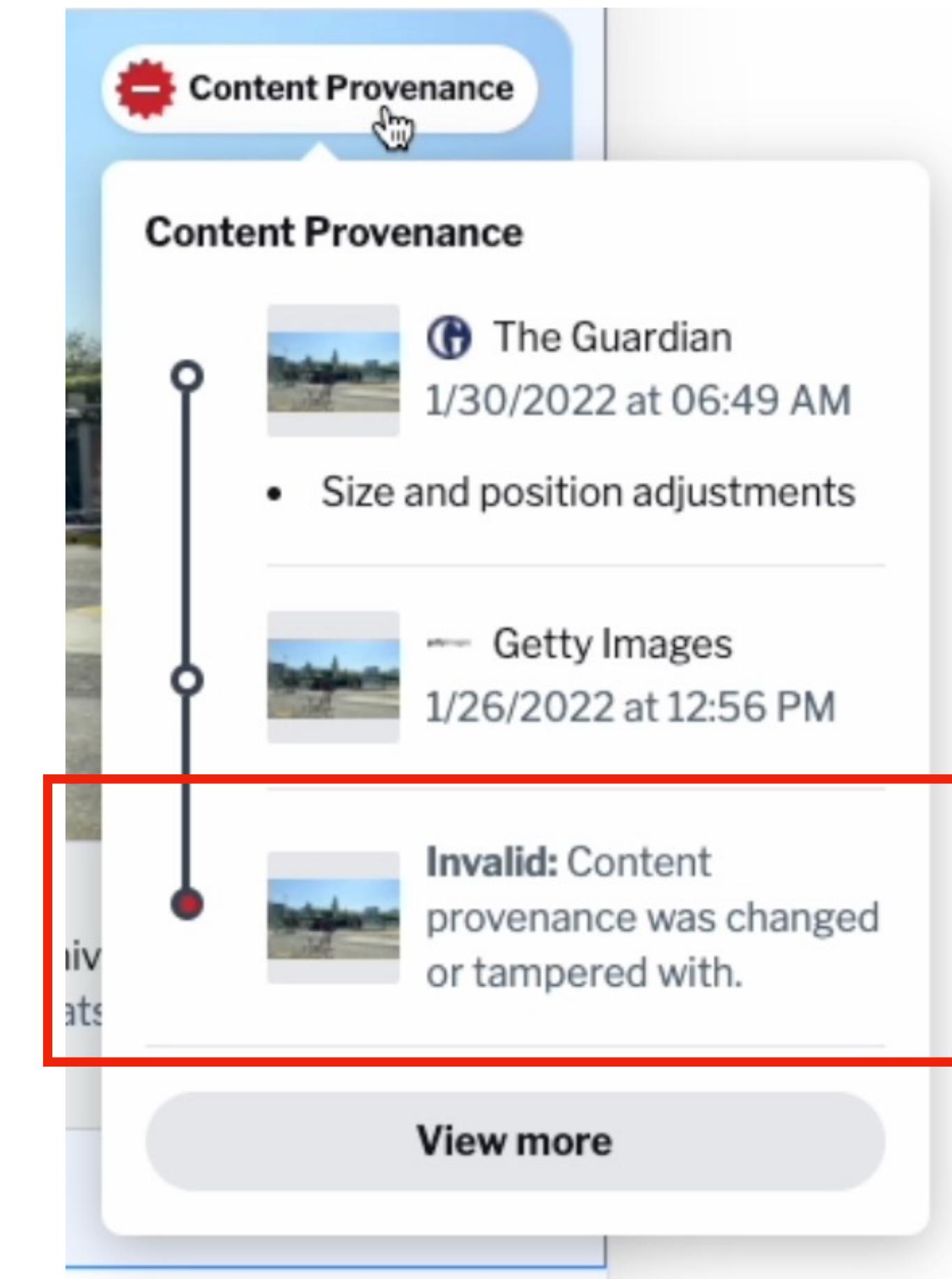


Introduction

Incomplete and Invalid Provenance Chains



Incomplete chain



Invalid chain

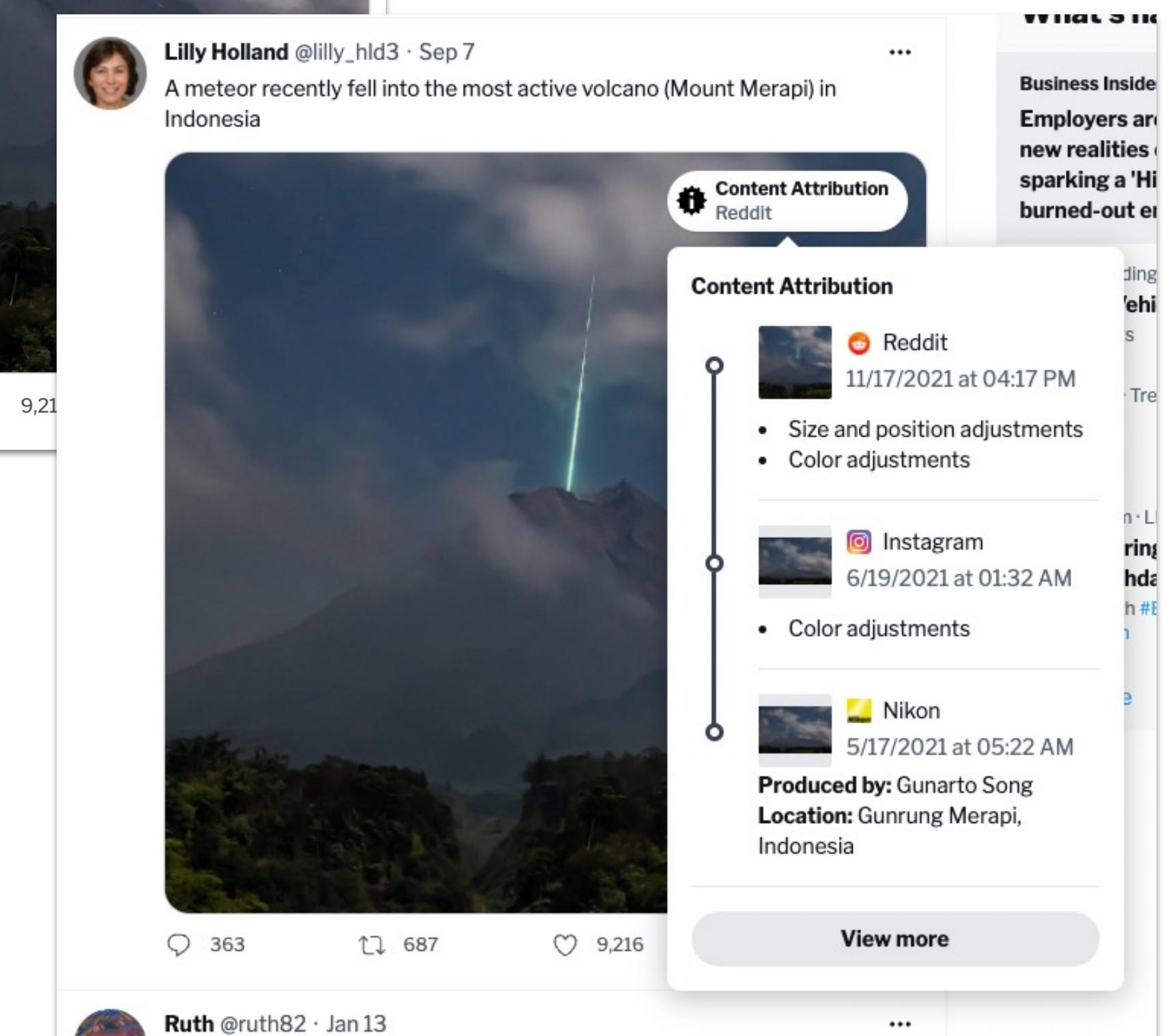
The goal of provenance is to...
***supply users with rich
information to empower them
to make better credibility
judgements.***

How does provenance information impact users' credibility judgements in social media feeds?

Overview

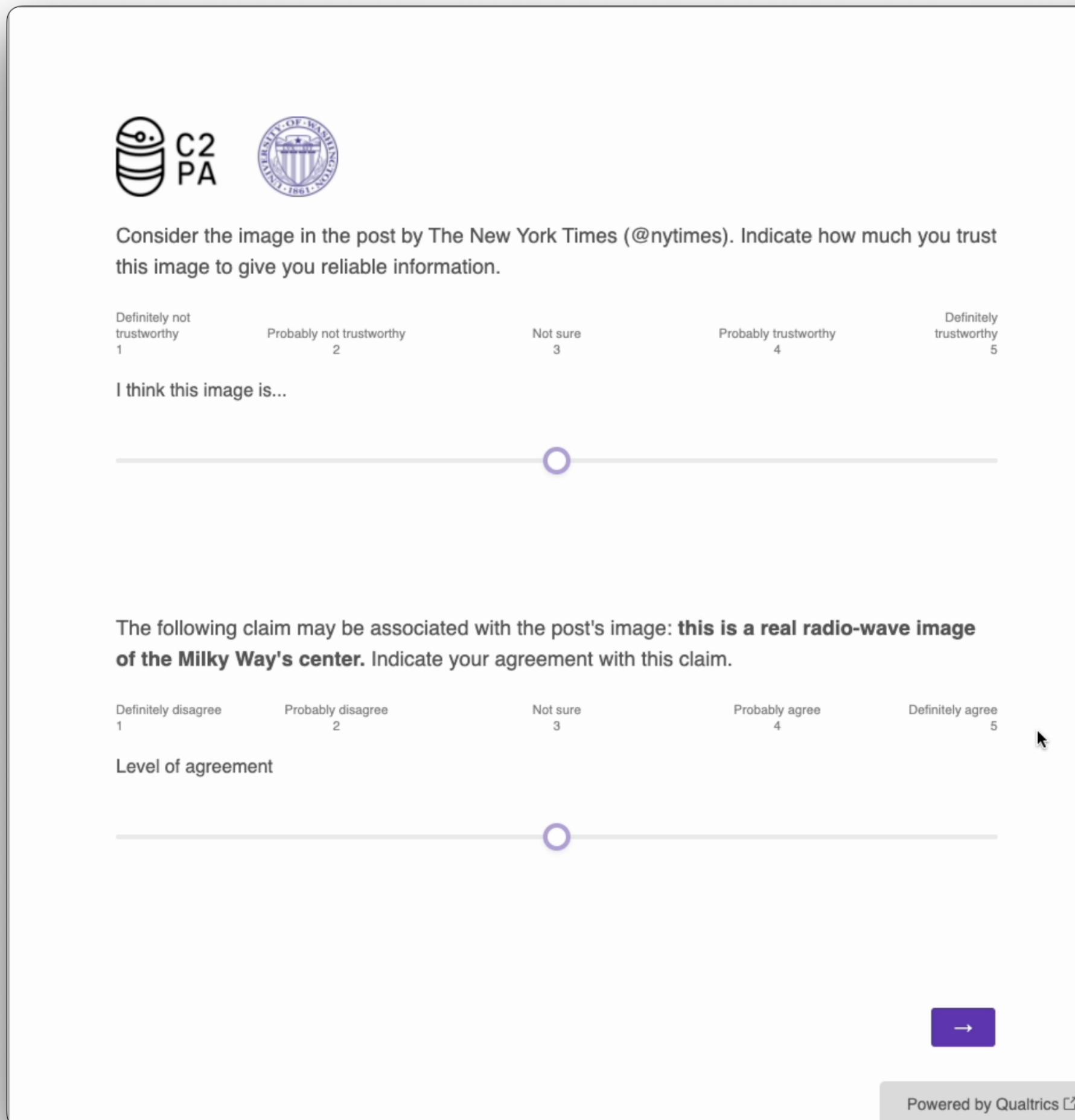
Experiment

- **595 participants** from Prolific
 - 298 from US, 297 from UK
- **2 rounds of tasks**
 - Round 1: interact with Twitter-like feed and answered questions
 - Round 2: interact with same feed but *with media provenance info* and answer same questions



Overview

Study Interface



The survey interface displays a question about the trustworthiness of an image from The New York Times. The question asks: "Consider the image in the post by The New York Times (@nytimes). Indicate how much you trust this image to give you reliable information." Below the question is a five-point Likert scale ranging from "Definitely not trustworthy" (1) to "Definitely trustworthy" (5). A slider bar allows the participant to select a value between 1 and 5. The C2 PA logo and the University of Washington seal are visible at the top left.

Consider the image in the post by The New York Times (@nytimes). Indicate how much you trust this image to give you reliable information.

Definitely not trustworthy 1 Probably not trustworthy 2 Not sure 3 Probably trustworthy 4 Definitely trustworthy 5

I think this image is...

—○—

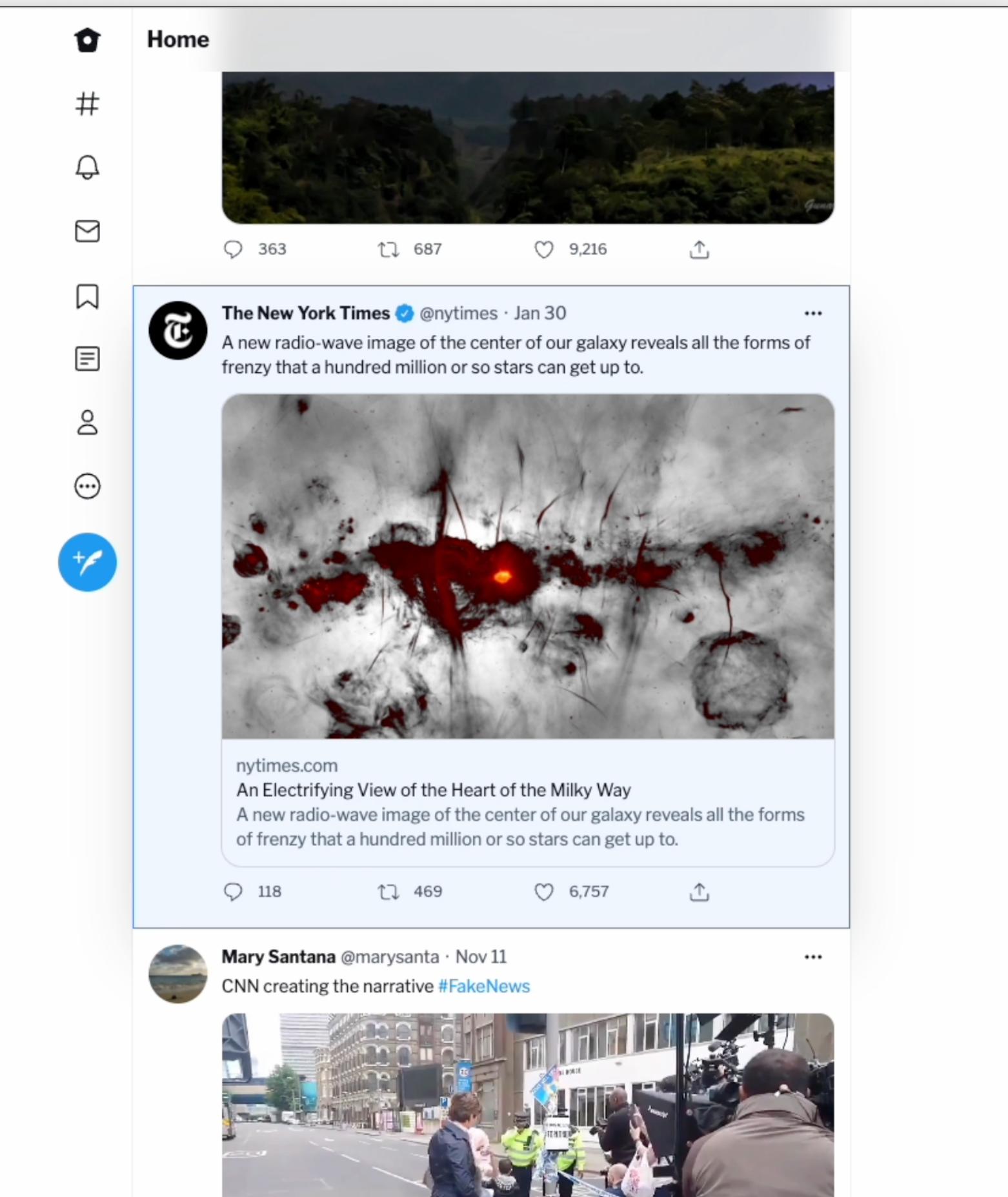
The following claim may be associated with the post's image: **this is a real radio-wave image of the Milky Way's center**. Indicate your agreement with this claim.

Definitely disagree 1 Probably disagree 2 Not sure 3 Probably agree 4 Definitely agree 5

Level of agreement

—○—

Powered by Qualtrics →



The social media prototype shows a feed of three posts. The first post is from The New York Times (@nytimes) on Jan 30, featuring a radio-wave image of the Milky Way's center. It has 363 comments, 687 shares, and 9,216 likes. The second post is from Mary Santana (@marysanta) on Nov 11, featuring a video of a CNN news crew. It has 118 comments, 469 shares, and 6,757 likes. The third post is a placeholder for a survey response.

Home

363 687 9,216

The New York Times @nytimes · Jan 30
A new radio-wave image of the center of our galaxy reveals all the forms of frenzy that a hundred million or so stars can get up to.

nytimes.com
An Electrifying View of the Heart of the Milky Way
A new radio-wave image of the center of our galaxy reveals all the forms of frenzy that a hundred million or so stars can get up to.

118 469 6,757

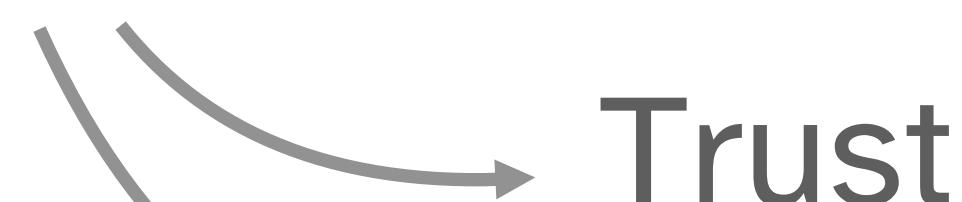
Mary Santana @marysanta · Nov 11
CNN creating the narrative #FakeNews

Survey

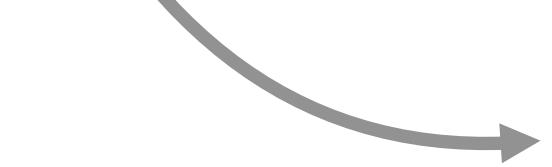
Social media prototype

Definitions

The goal of provenance is to...
***supply users with rich
information to empower them
to make better credibility
judgements.***



Trust



Perceived Accuracy

Definitions

Trust

Consider the image in the post by The Guardian (@guardian). Indicate how much you trust this image to give you reliable information.

Definitely not
trustworthy
1

Probably not trustworthy
2

Not sure
3

Probably trustworthy
4

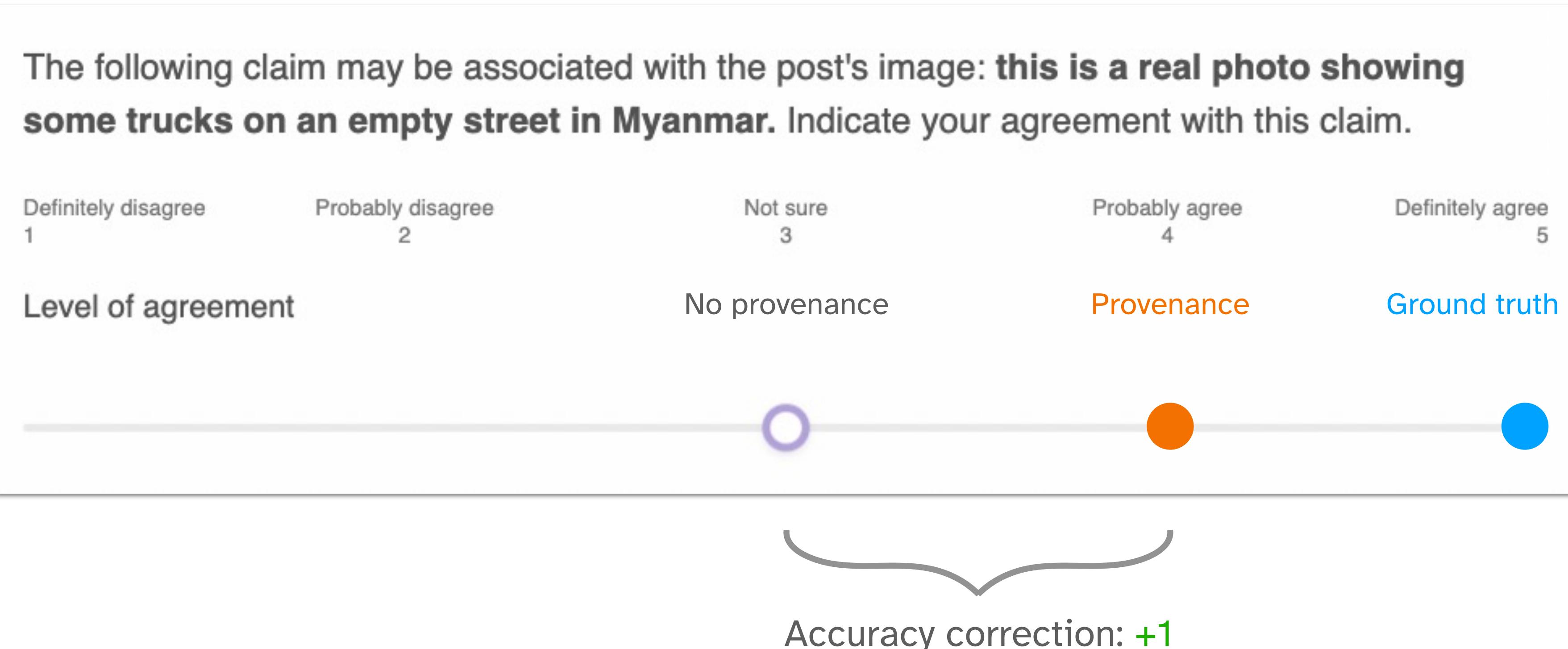
Definitely
trustworthy
5

I think this image is...



Definitions

Perceived Accuracy



Definitions

Normal, Incomplete, Invalid Provenance



Normal

No abnormalities with provenance info

1/2 of participants



Incomplete

Media changed w/o using provenance-enabled tool

1/2 of participants



Invalid

Provenance info has been tampered with

Findings

1.

Perceptions of incomplete or invalid provenance info

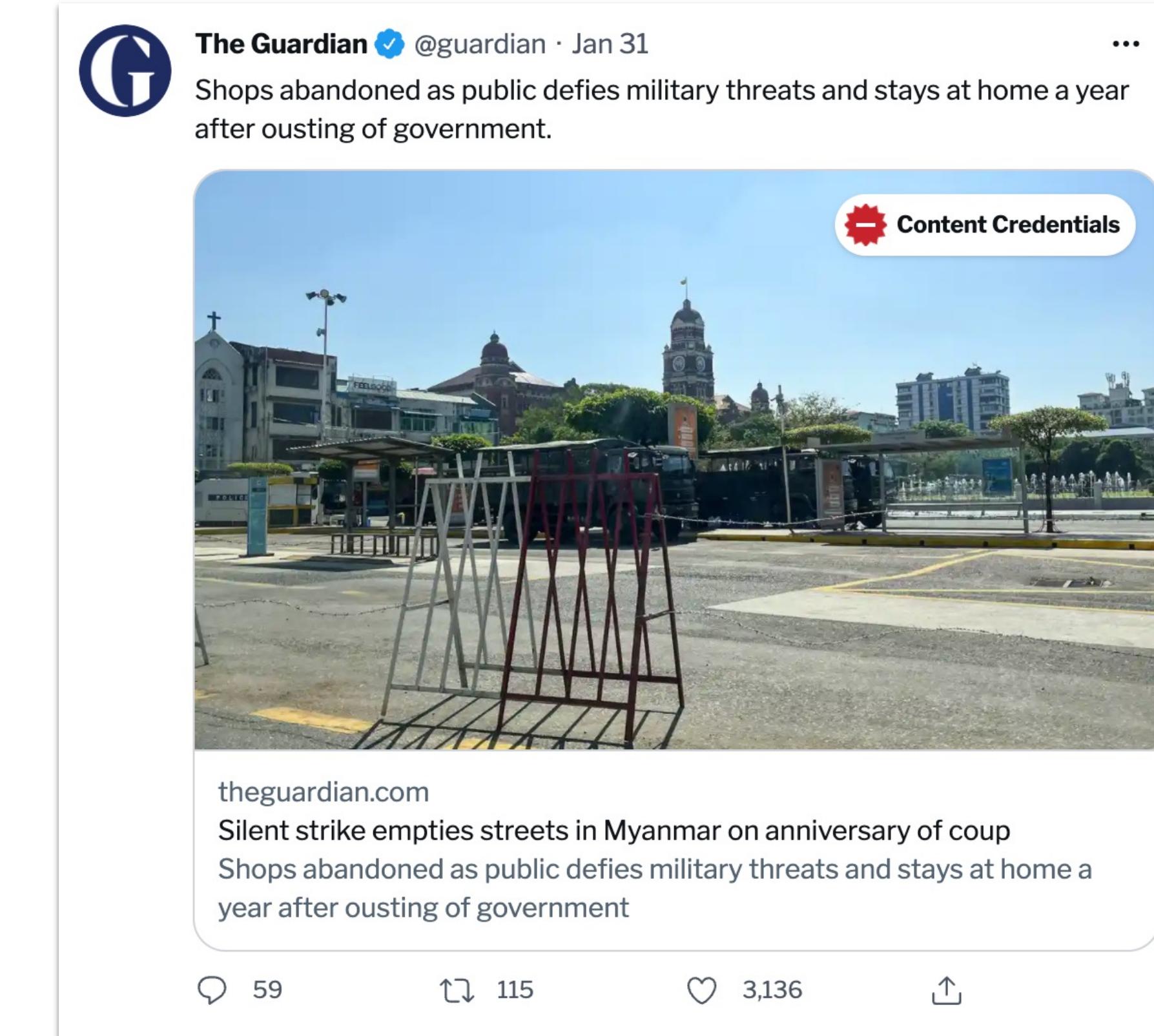
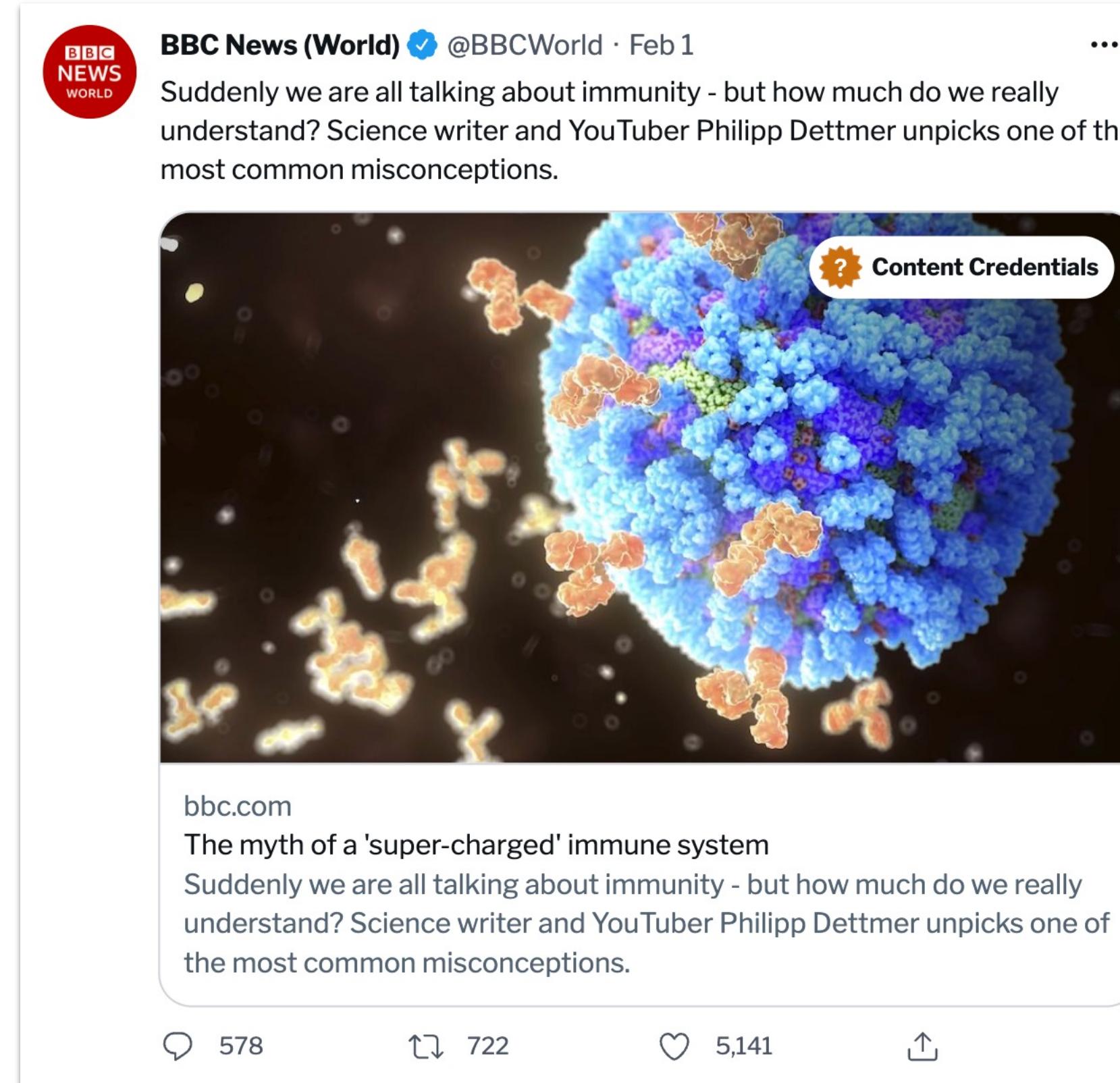
2.

Changes in trust and accuracy perceptions post-provenance

3.

Users' comprehension of provenance

Perceptions of incomplete or invalid provenance information Incompleteness and Invalidity Can Worsen Correction



Normal correction: **-0.02*****
Incomplete correction: **-0.43*****

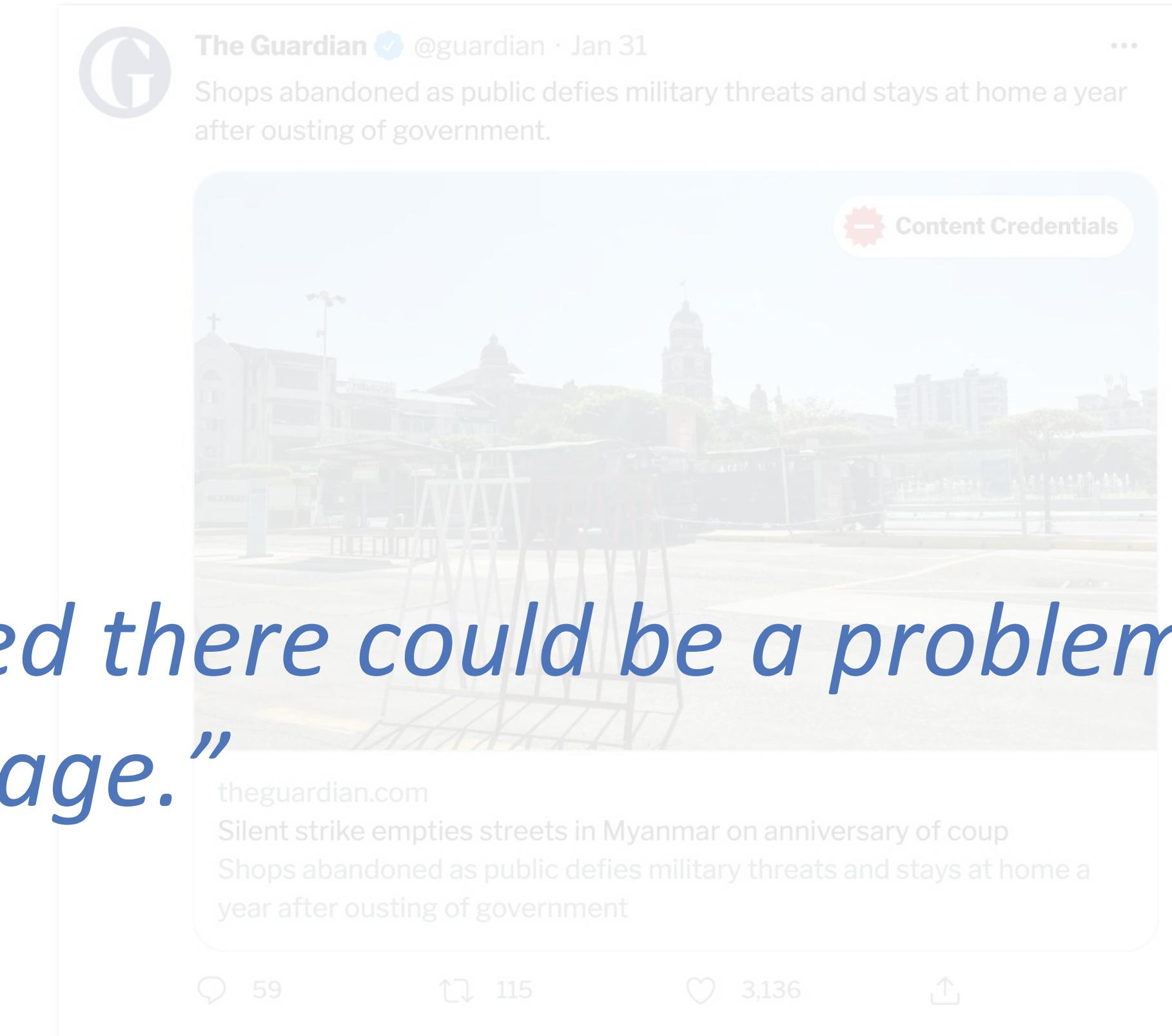
$p < 0.05^*$, $p < 0.01^{**}$, $p < 0.001^{***}$

Normal correction: **-0.09*****
Invalid correction: **-1.16*****

Perceptions of incomplete or invalid provenance information Incompleteness and Invalidity Can Worsen Correction



“I like that the [invalidity] indicated there could be a problem with the image.”



Normal correction: -0.02***
Incomplete correction: -0.43***

$p < 0.05^*$, $p < 0.01^{**}$, $p < 0.001^{***}$

Normal correction: -0.09***
Invalid correction: -1.16***

Findings

1.

Perceptions of incomplete or invalid provenance info

2.

Changes in trust and accuracy perceptions post-provenance

3.

Users' comprehension of provenance

Changes in trust and accuracy perceptions post-provenance Provenance Works! Sometimes

For *non-deceptive* content:

	Δ trust	Correction
Normal	+0.19***	+0.20***
Incomplete	-0.53***	-0.45***
Invalid	-1.14***	-1.24***

For *deceptive* content:

	Δ trust	Correction
Normal	-0.15**	+0.16*
Incomplete	-0.66***	+0.86***
Invalid	-0.68***	+0.57***

$p < 0.05^*$, $p < 0.01^{**}$, $p < 0.001^{***}$

Findings

1.

Perceptions of incomplete or invalid provenance info

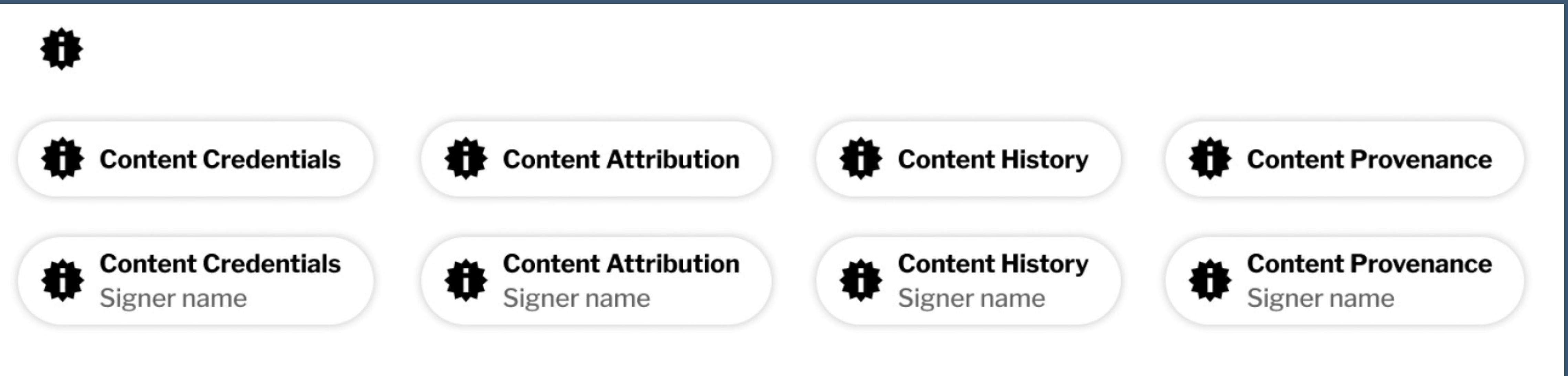
2.

Changes in trust and accuracy perceptions post-provenance

3.

Users' comprehension of provenance

Icon variation



Language variation

Users' understanding of provenance

No significant differences. However...

Valuable insights from qualitative responses:

"Some people might not be familiar with the word 'provenance.'

'Origin' might be more familiar.

*"A pop up **definition for what some of the terms mean** would be helpful."*

*"Better icons would be great; **green for trustworthy, red for untrustworthy, grey for unsure.**"*

*"Perhaps a **more simple colour coded green, amber, red system of authenticity.**"*

Summary & Implications

Summary & Implications

Towards Usable Provenance

Rethinking incomplete and invalid provenance

Can they be used adversarially?

Trading off quick judgement and investigative reasoning

It's hard to slow down on social media, but provenance requires users to carefully reason about their content.

Not all edits are created equal

Many of our heavier edits were deceptive in nature. What about more innocent edits?

Summary & Implications

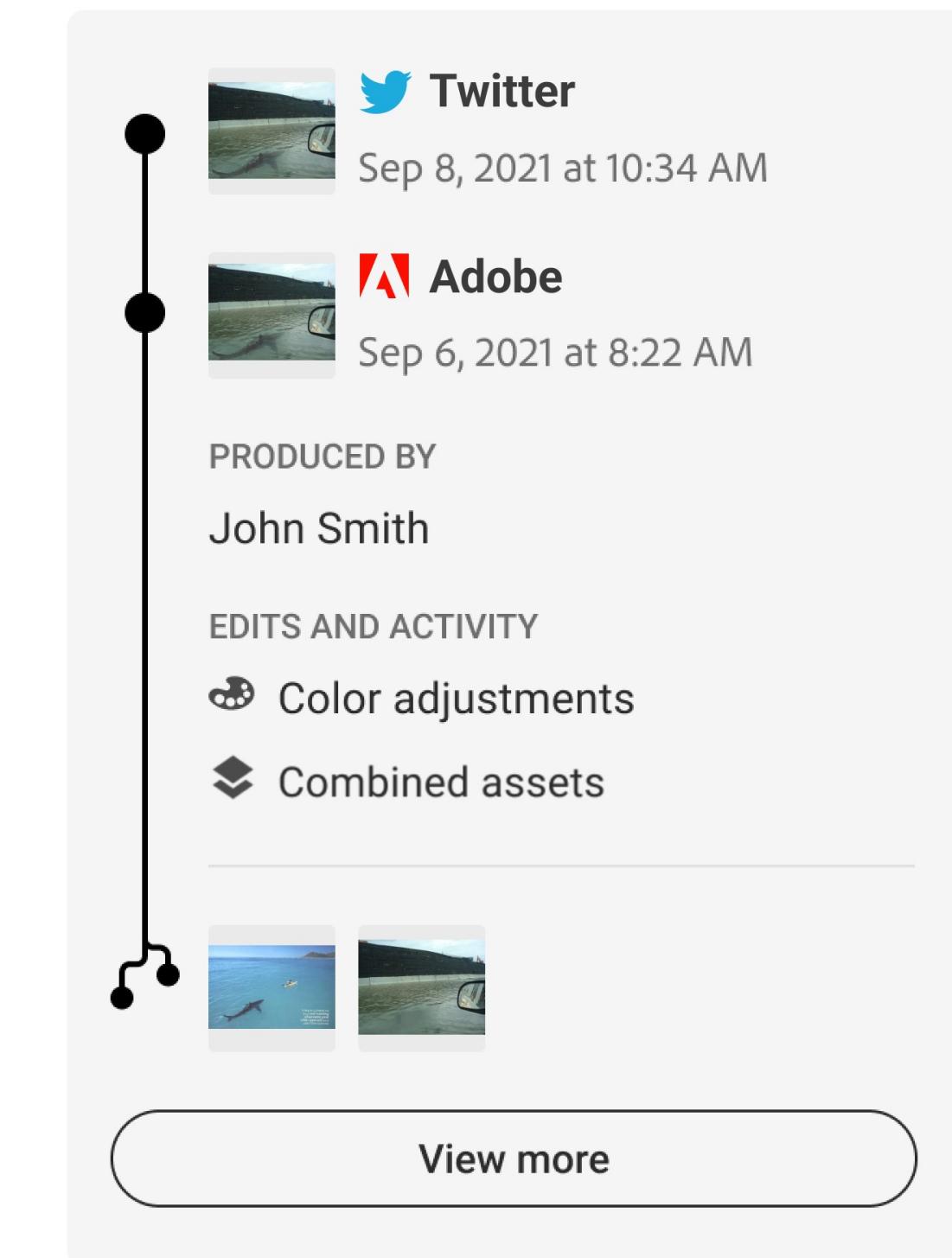
Impact & Generative AI

Generative AI disclosures

Integrated into Adobe Firefly. Empowers artists to get credit for their work.

UI and iconography redesign

Incomplete and invalid provenance has been redesigned.



2022

2023

Content Credentials

Issued by Adobe Inc. on Oct 5, 2023

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

Produced by obidigbo nzeribe

Social media [Instagram](#), [Behance](#)

App or device used

Adobe Photoshop 25.1.0

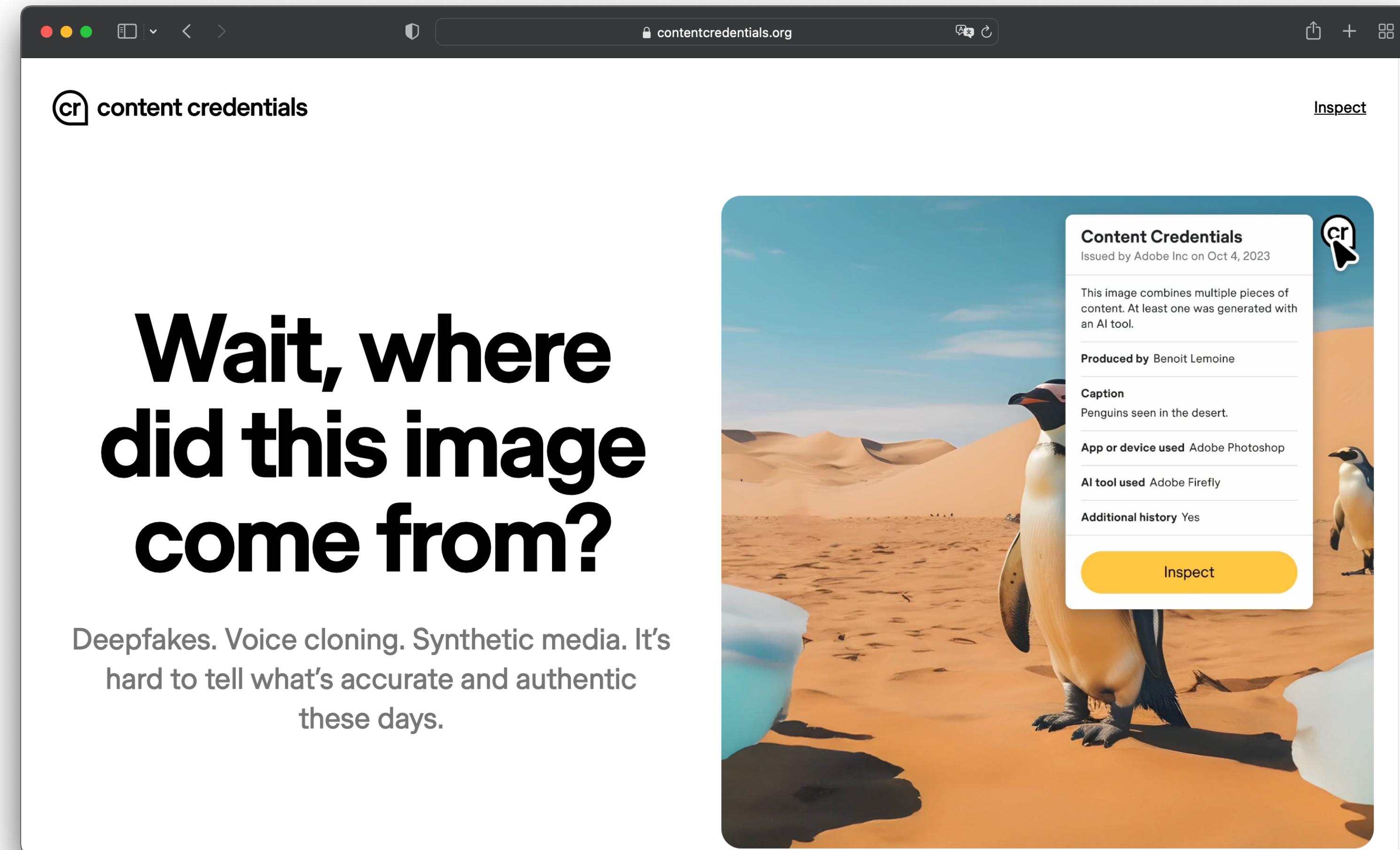
AI tool used Adobe Firefly

Additional history Yes

Inspect

Summary & Implications

Check Out Content Credentials!



<https://contentcredentials.org>

Thanks!

Examining the Impact of Provenance-Enabled Media on Trust and Accuracy Perceptions



K. J. Kevin Feng, University of Washington
Nick Ritchie, BBC UX&D
Pia Blumenthal, Adobe
Andy Parsons, Adobe
Amy X. Zhang, University of Washington

Contact: kjfeng@uw.edu | kjfeng.me

CSCW 2023
October 14-18, 2023 | Minneapolis, MN

Scan for paper:



*Special thanks to the
C2PA Technical
Working Group and
C2PA UX Taskforce
for their support and
feedback!*



Social
Futures
Lab

DUB
DESIGN
USE
BUILD

W C2
PA