

Dopad Generative AI a Deepfake na transparentnost a autenticitu informací

A jaké máme možnosti zajistit jejich původ a integritu

1.6.2023

AFCEA – BS70 – Bezpečnostní seminář:
Kybernetická bezpečnost & umělá inteligence

Roman Cinkais



3KEYCOMPANY



GENERATIVE AI

Generativní umělá inteligence (neboli Generative AI) je typ umělé inteligence (AI) schopný generovat text, obrázky nebo jiná média v reakci na různé vstupy a podněty. Generativní modely umělé inteligence se učí vzorce (tzv. patterny) a strukturu svých vstupních dat a poté generují nová data, která mají podobné vlastnosti, nicméně jsou uměle vytvořena.

DALL-E 2

DALL-E 2 is an AI system that can create realistic images and art from a description in natural language.

Introducing ChatGPT

We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests.

Your AI pair programmer

GitHub Copilot uses the OpenAI Codex to suggest code and entire functions in real-time, right from your editor.

FEATURE AI-POWERED

Bing Chat

With the new Bing built into the Microsoft Edge sidebar, you can ask complex questions, find comprehensive answers, get summarized information, find inspiration to build upon all in a side-by-side view, with no need to flip between tabs.

Create Presentation Slides with AI in seconds

Say goodbye to tedious, manual slides creation. Let the AI write the outline and presentation content for you. With our tool, you can easily create professional, engaging slides from any text in no time.

Bard can write some lyrics for your heartbreak anthem titled "Lovesick"

Meet Bard: your creative and helpful collaborator, here to supercharge your imagination, boost your productivity, and bring your ideas to life.

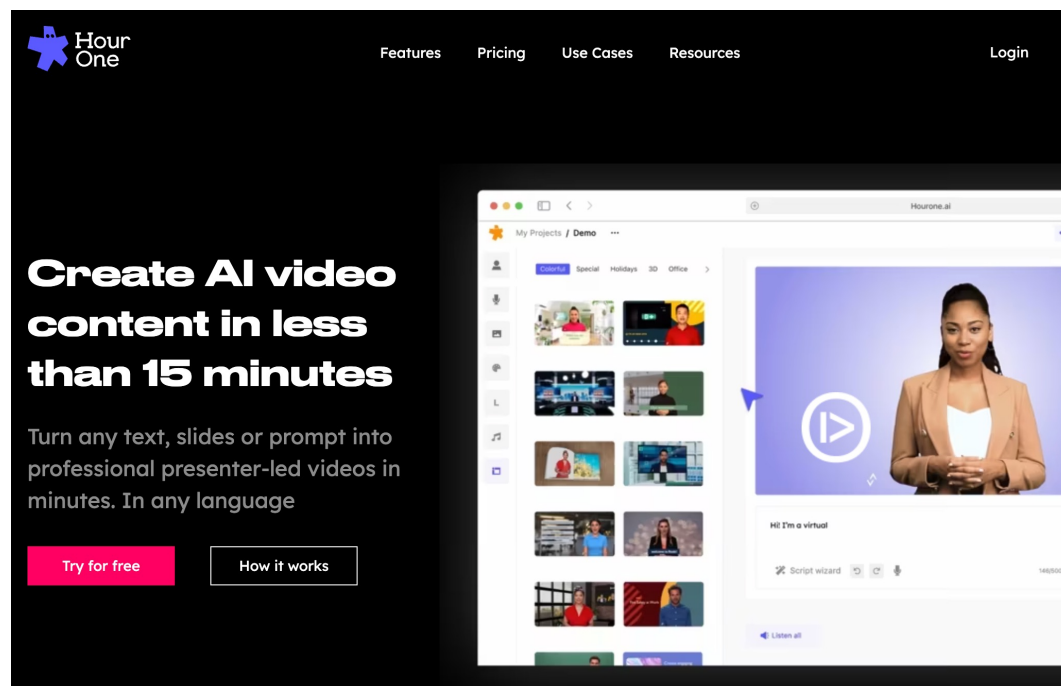
Bard isn't currently supported in your country. Stay tuned!

Market Map: Generative AI for Virtual Worlds

Experience	Discovery	Creator Economy	Spatial Computing	Decentralize	Human Interface	Infrastructure
OpenAI PRISMA LABS hidden door EPIC GAMES Character.AI Replika latitude amazon Google FABLE REGRESSION GAMES UNLEASHED THE CULTURE DAO ONEIROCOM SYSTEMS SPELLBRUSH CATHEDRAL STUDIOS Meta	Meta Google Hugging Face Riku kaggle Microsoft Perplexity PromptBase Microsoft	OpenAI Midjourney Figma replit scenario supertone runway neosapience GitHub Copilot METAPHYSIC Replicate rct.ai boomy Leonardo.AI CEPETTO DAACI modl.ai F C LOVELLACE STUDIOS READY PLAYER ME SPLASH USICO inworld LAIKA charisma.ai Aflorithmic convai Inpris AUTODESK REVIT	NVIDIA OpenAI Google AI Meta LUMA AI Sloyd KAEDIM OPUS plask Kinetix A armory MOVE DEEPMOTION RADICAL Unity EPIC GAMES MASTERPIECE STUDIO	stability.ai Hugging Face deepset NVIDIA THE LINUX FOUNDATION Google Microsoft Meta	OpenAI Meta Google AI21labs NVIDIA EleutherAI co:here Adept amazon Apple personal.ai	NVIDIA Apple AMD Qualcomm SAMSUNG aws ASML Google Microsoft Me

DEEPPFAKE

Deepfake jsou syntetická média, která byla digitálně zmanipulována tak, aby přesvědčivě nahradila podobu jedné osoby podobou jiné osoby, virtuálně. I když akt vytváření falešného je obecně znám dlouhou dobu, deepfake využívá strojové učení a umělou inteligence k manipulaci nebo generování vizuálního a zvukového obsahu, který může snadno klamat, protože vypadá reálně.



synthesia

Create a free AI video

Select a template & edit your script. Political, sexual and discriminatory content will not be approved.

1 SELECT VIDEO TEMPLATE

Synthesia Demo **Sales Pitch** Learning & Development
Compliment How-To Video

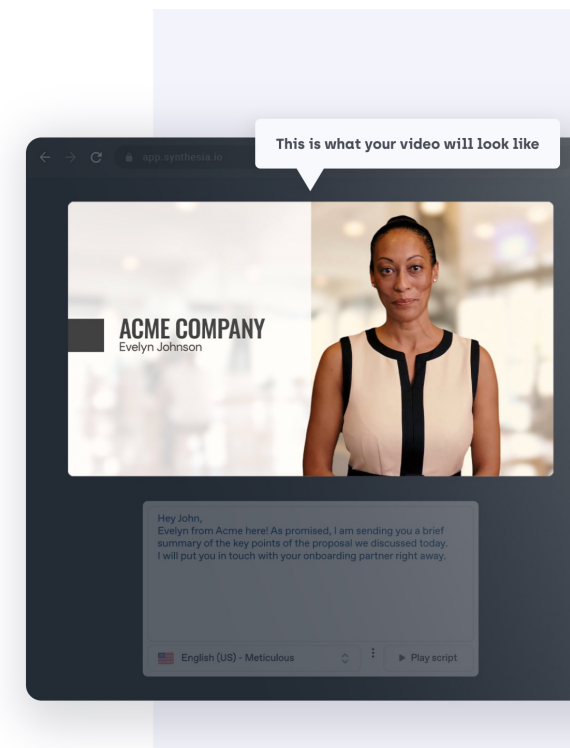
2 EDIT YOUR VIDEO SCRIPT

You can use any popular language

Hey John, Evelyn from Acme here! As promised, I am sending you a brief summary of the key points of the proposal we discussed today. I will put you in touch with your onboarding partner right away.

3 characters left

Continue



VLASTNICTVÍ

Vztahuje se porušení autorských práv, patentů nebo ochranných známek na generovaný obsah?

Je jasné, kdo vlastní obsah, který vytvořila umělá inteligence pro zákazníka?

Kam generativní AI zapadá z právního pohledu?

Byla umělá inteligence trénována na licencovaném obsahu?

Mají být uživatelé AI odpovědní za porušování autorských práv a ochranných známek?

Může vývojář použít zdrojový kód generovaný AI bez ohledu na vlastnické práva?

Pokud díla vytvořená generativním AI nejsou způsobilá pro autorská práva, jaký je jejich právní status?

Může generovaný obsah někdo vlastnit – a pokud ano, kdo?

VODOZNAK A METADATA



Digitální vodoznak poskytuje alternativní řešení pro zajištění odolnosti proti neoprávněné manipulaci, duševní vlastnictví a posílení bezpečnosti multimediálních dokumentů. Jakýkoli digitální obsah, jako jsou obrázky, zvuk a videa, a další, může obsahovat vodoznak.



Digitální steganografie jako protipól digitálního vodoznaku se zabývá způsoby, jak vložit informaci do multimediálního obsahu takovým způsobem, že přítomnost informace není pro lidi jednoduše identifikovatelná (viditelná). Někdy se označuje taky jako krytá komunikace.



Exif (zkratka z anglického Exchangeable image file format) je specifikace pro formát metadat, vkládaných do souborů digitálními fotoaparáty (včetně chytrých mobilních telefonů), skenery a dalšími zařízeními nebo programy pro zpracování obrázků, případně též souborů se záznamem zvuku.

Dopad Generative AI a Deepfake na transp a autenticitu informací

A jaké máme možnosti zajistit jejich původ a integritu

1.6.2023

AFCEA – BS70 – Bezpečnostní seminář:
Kybernetická bezpečnost & umělá inteligence

Roman Cinkais

3KEYC

```
BS70 – Bezpečnostní seminář - Kybernetická bezpečnost a umělá inteligence — -zsh — 93...
ExifTool Version Number      : 12.60
File Name                    : 2306_3Key_AFCEA_BS70_2023_watermark.jpg
Directory                    : .
File Size                    : 177 kB
File Modification Date/Time   : 2023:05:29 15:55:08+02:00
File Access Date/Time        : 2023:05:29 15:56:19+02:00
File Inode Change Date/Time   : 2023:05:29 15:56:17+02:00
File Permissions              : -rw-r--r--
File Type                    : JPEG
File Type Extension          : jpg
MIME Type                    : image/jpeg
Photoshop Quality             : 10
Photoshop Format              : Progressive
Progressive Scans             : 3 Scans
Current IPTC Digest           : cdcffa7da8c7be09057076aeaf05c34e
Coded Character Set           : UTF8
Application Record Version    : 0
IPTC Digest                   : cdcffa7da8c7be09057076aeaf05c34e
XMP Toolkit                   : Adobe XMP Core 9.0-c001 79.c0204b2def, 2023/02/02-12:14:24
Original Document ID          : 3116821B24F0248FF4574CAFDCB119FE
Document ID                   : xmp.did:3BC526DA5C63834EBAE14CD3E33AE3C1
Instance ID                   : xmp.iid:E0AA678370C19546858E02219D5222B5
ICC Profile Name              : sRGB IEC61966-2.1
Creator Tool                  : Adobe Photoshop 24.5 (Macintosh)
Derived From Instance ID      : xmp.iid:dfef11a1c-dc32-4b25-97cc-2a1ac85e7409
Derived From Document ID      : adobe:docid:photoshop:61694d11-1ce9-064b-9b86-6b58d3871db4
Profile CMM Type              : Linotronic
Profile Version               : 2.1.0
Profile Class                 : Display Device Profile
Color Space Data              : RGB
Profile Connection Space      : XYZ
Profile Date Time             : 1998:02:09 06:49:00
Profile File Signature        : acsp
Primary Platform              : Microsoft Corporation
CMM Flags                     : Not Embedded, Independent
Device Manufacturer           : Hewlett-Packard
Device Model                  : sRGB
Device Attributes              : Reflective, Glossy, Positive, Color
Rendering Intent               : Perceptual
Connection Space Illuminant   : 0.9642 1 0.82491
Profile Creator                : Hewlett-Packard
Profile ID                    : 0
Profile Copyright              : Copyright (c) 1998 Hewlett-Packard Company
Profile Description            : sRGB IEC61966-2.1
Media White Point              : 0.95045 1 1.08905
Media Black Point             : 0 0 0
Red Matrix Column              : 0.43607 0.22249 0.01392
Green Matrix Column            : 0.38515 0.71687 0.09708
Blue Matrix Column             : 0.14307 0.06061 0.7141
Device Mfg Desc                : IEC http://www.iec.ch
Device Model Desc              : IEC 61966-2.1 Default RGB colour space - sRGB
Viewing Cond Desc              : Reference Viewing Condition in IEC61966-2.1
Viewing Cond Illuminant       : 19.6445 20.3718 16.8089
Viewing Cond Surround         : 3.92889 4.07439 3.36179
Viewing Cond Illuminant Type   : D50
Luminance                     : 76.03647 80 87.12462
```


DIGITÁLNÍ PODEPISOVÁNÍ

eIDAS - zaručený elektronický podpis musí splňovat tyto požadavky:

- je jednoznačně spojen s podepisující osobou;
- umožňuje identifikaci podepisující osoby;
- je vytvořen pomocí dat pro vytváření elektronických podpisů, která podepisující osoba může s vysokou úrovní důvěry použít pod svou výhradní kontrolou; a
- je k datům, která jsou tímto podpisem podepsána, připojen takovým způsobem, že je možné zjistit jakoukoliv následnou změnu dat.

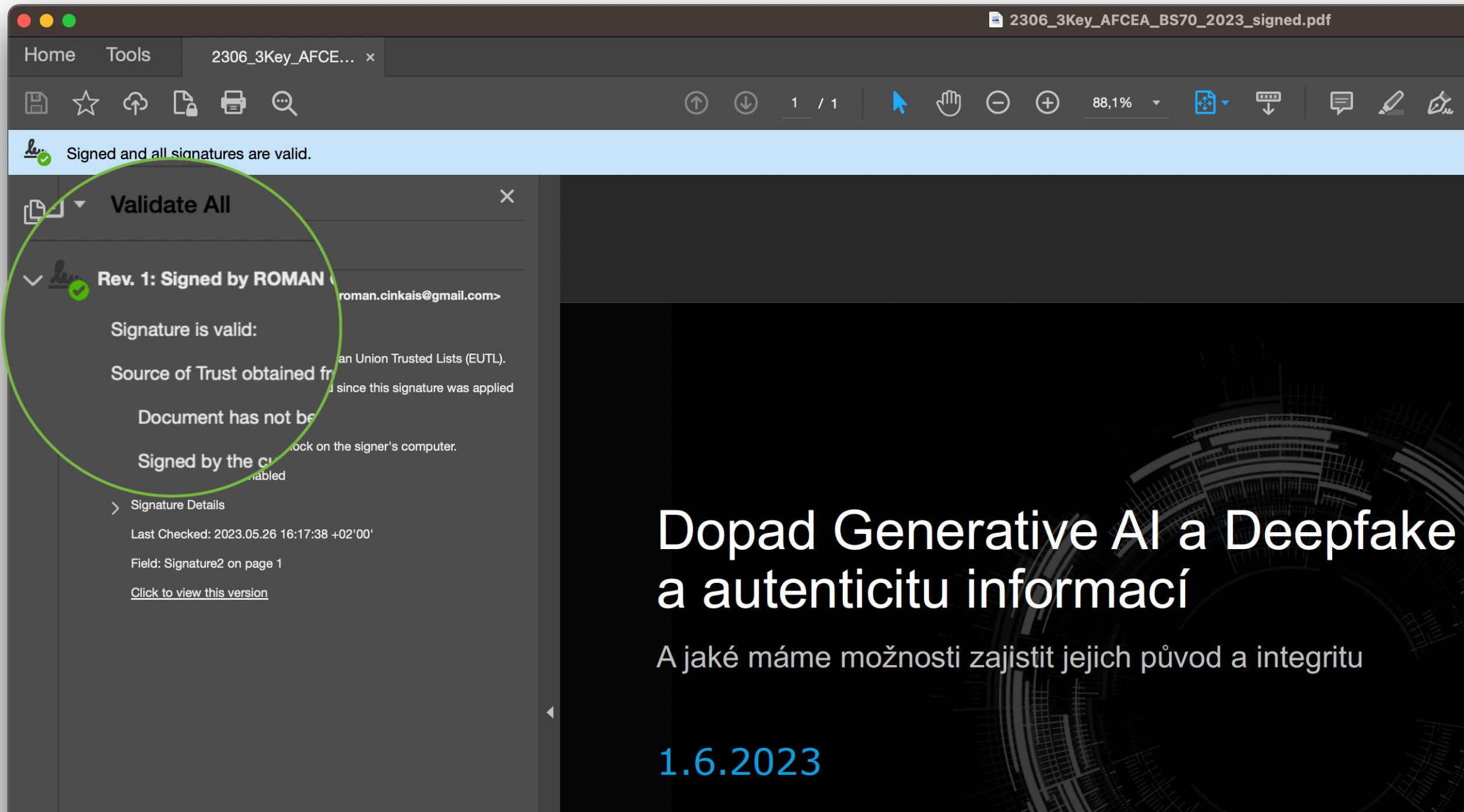


STANDARDY:

ETSI EN 319 142 - PDF Advanced Electronic Signature Profiles (PAdES)
ETSI EN 319 132 - XML Advanced Electronic Signatures (XAdES)
ETSI EN 319 122 - CMS Advanced Electronic Signatures (CAdES)
ETSI TS 119 182-1 (draft) JSON Advanced Electronic Signatures (JAdES)
ETSI EN 319 162 - Associated Signature Containers (ASiC)
ETSI EN 319 422 - Time-stamping protocol and time-stamp token profiles

PODEPISOVÁNÍ:

- Dokumenty a smlouvy
- Transakce
- Zdrojový kód a instalační balíčky
- Firmware
- Čas (časové razítko)
- Pasy a občanky
- Datové schránky
- Ochranná známka



 Signed and all signatures are valid.

 **Validate All** 

 **Rev. 1: Signed by ROMAN CINKAIS (roman.cinkais@gmail.com)**

Signature is valid:

Source of Trust obtained from an Union Trusted Lists (EUTL).

Document has not been modified since this signature was applied

Document has not been modified since this signature was applied

Signed by the clock on the signer's computer.

enabled

> Signature Details

Last Checked: 2023.05.26 16:17:38 +02'00'

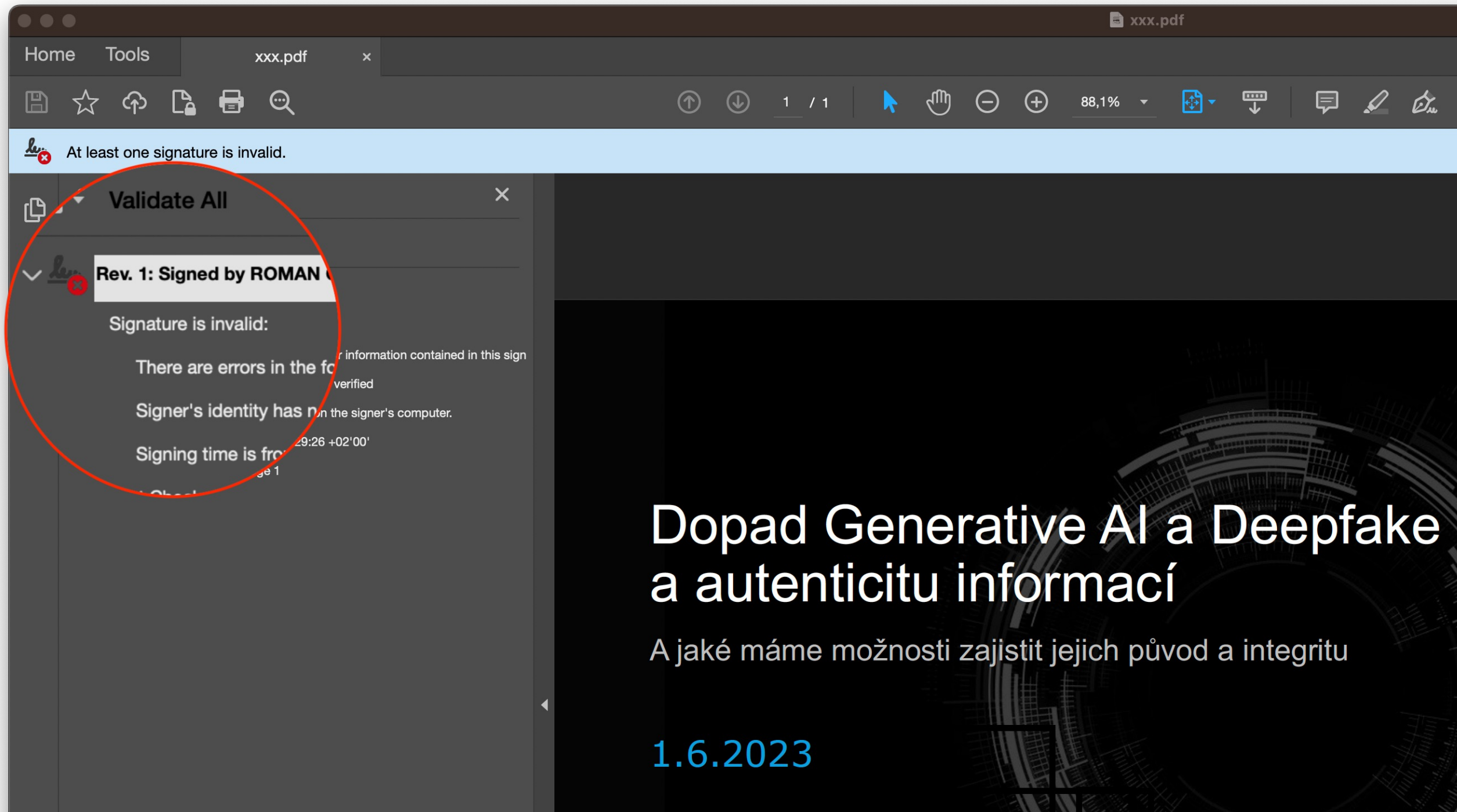
Field: Signature2 on page 1

[Click to view this version](#)

Dopad Generative AI a Deepfake a autenticitu informací

A jaké máme možnosti zajistit jejich původ a integritu

1.6.2023



Home Tools

xxx.pdf



88,1%



At least one signature is invalid.



Validate All



Rev. 1: Signed by ROMAN C

Signature is invalid:

There are errors in the following information contained in this signature:
The signature cannot be verified

Signer's identity has not been verified on the signer's computer.

Signing time is from 2023-09-26 +02'00'

Dopad Generative AI a Deepfake a autenticitu informací

A jaké máme možnosti zajistit jejich původ a integritu

1.6.2023

CODE SIGNING

- Podepisování kódu je proces digitálního podepisování dat potvrzení autora softwaru a zaručení, že kód nebyl od podepsání změněn.
- Definice „kódu“ se mění. S příchodem moderních technologií a postupů, jako je jsou například softwarově definovaná infrastruktura, nebo agilní DevOps přístup, podepisujeme nejen softwarový kód.
- Prakticky každá společnost podepisuje kód v nějaké podobě nebo formě.
- Podepisování kódu je součástí bezpečného vývoje softwaru. Automatizace pomáhá pro zvýšení logické bezpečnosti.



Software

Artefakty

Kontejnery

Firmware

Skripty

Makra

Git

OCHRANA ZNAČEK (VMC)

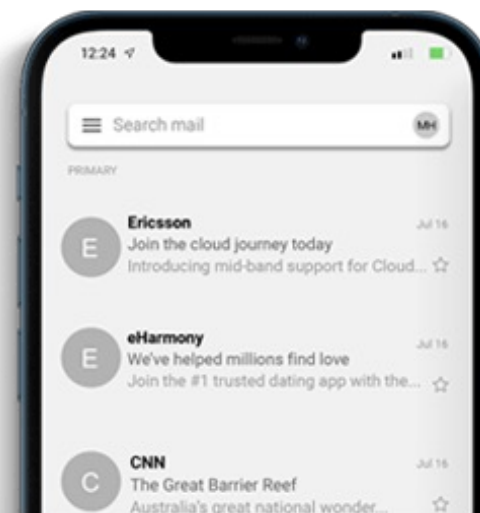
- Verified Mark Certificates (VMC)
- VMC představuje kryptograficky ověřitelnou a auditovatelnou vazbu mezi identitou, logem a doménou.
- Pomáhá chránit zákazníky a značku před phishingovými a spoofingovými útoky.
- Umožňuje organizacím zobrazovat logo registrované ochranné známky u odchozích e-mailů.

DMARC - Domain-based Message Authentication, Reporting and Conformance

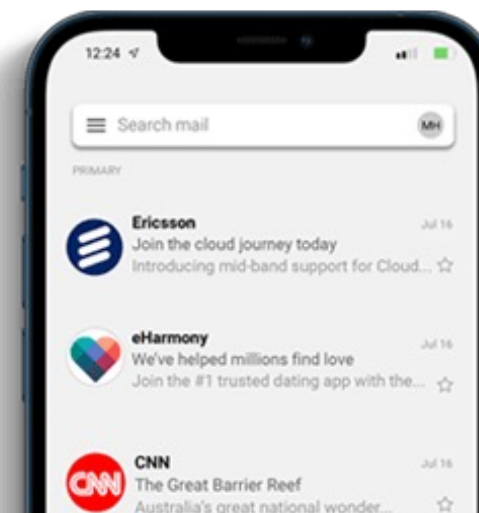
BIMI - Brand Indicators for Message Identification

```
v=BIMI1;l=https://www.3key.company/brand/bimi-  
logo.svg;a=https://www.3key.company/brand/certificate.  
pem
```

Without VMC



With VMC



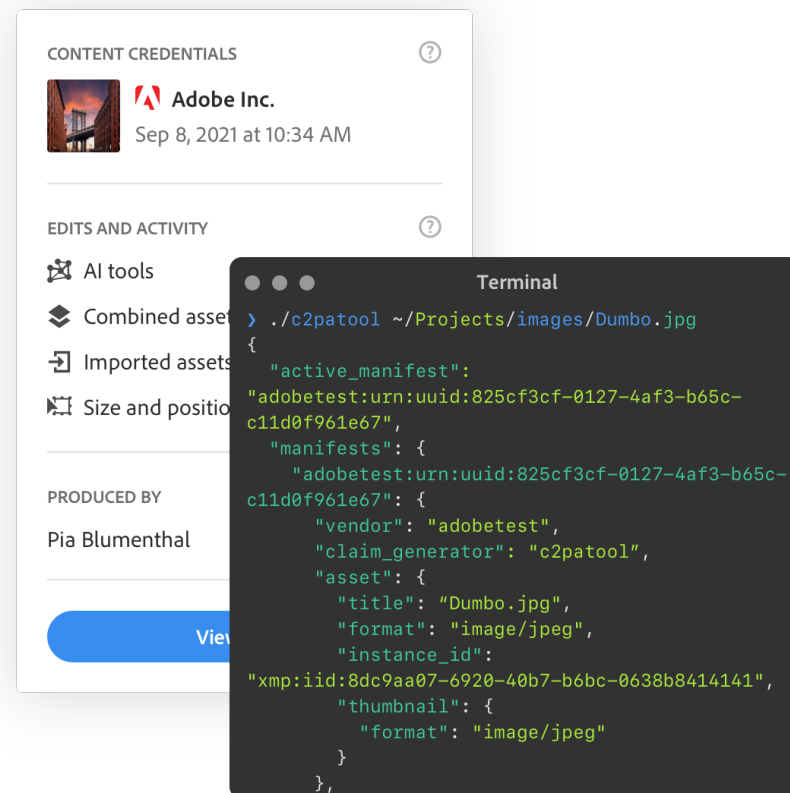
C2PA

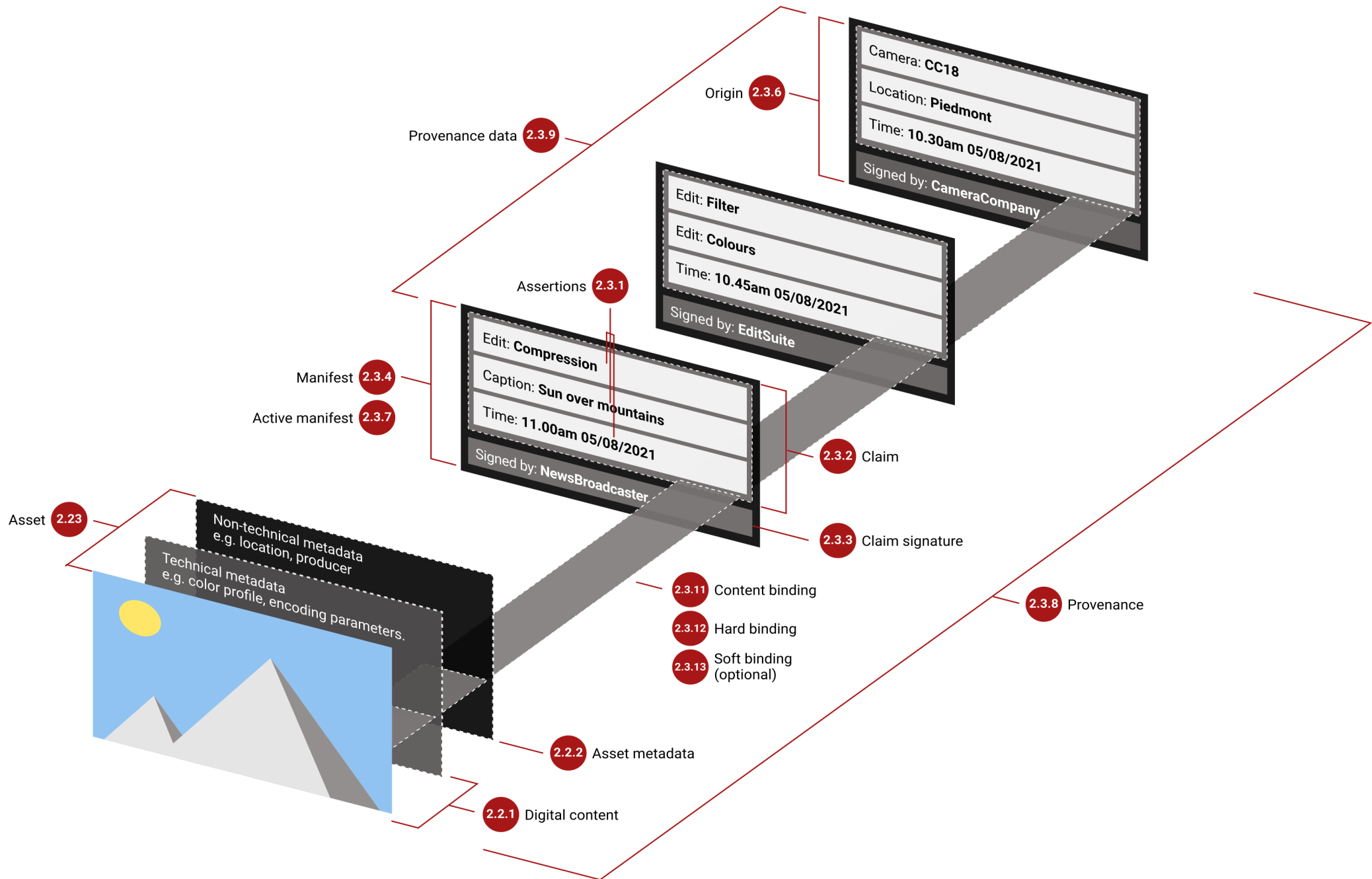


Coalition for Content Provenance and Authenticity (C2PA)

řeší rozšíření online zavádějících informací prostřednictvím vývoje technických standardů pro certifikaci zdroje a historie (a taky původu) mediálního obsahu.

C2PA sjednocuje úsilí iniciativy Adobe **Content Authenticity Initiative (CAI)**, která se zaměřuje na systémy poskytující kontext a historii digitálním médiím, a **Project Origin**, iniciativy pod vedením Microsoftu a BBC, která se zabývá dezinformacemi v ekosystému digitálních zpráv.





Content Credentials



2304_3Key_AFCEA_BS70_2023_dalibor_

SIGNED BY

 Adobe Inc.

May 26, 2023 at 3:39 PM GMT+2

PRODUCED BY

Roman Cinkais

PRODUCED WITH

 Adobe Photoshop 24.5.0

EDITS AND ACTIVITY

 Color or exposure edits

Adjusted properties like tone, saturation, curves, shadows, or highlights

 Combined

Composited or reordered content, or edited with content-sampling tools

 Drawing edits

Used tools like pencils, brushes, erasers, or shape, path, or pen tools

 Imported

```
c2patool — -zsh — 93x56
{
  "active_manifest": "adobe:urn:uuid:89969422-4c2b-4ff7-ac09-6b7fc148e7d3",
  "manifests": {
    "adobe:urn:uuid:89969422-4c2b-4ff7-ac09-6b7fc148e7d3": {
      "claim_generator": "Adobe_Photoshop/24.5.0 (build 20230512.r.500 3fc8e4c; mac) cai-help
er/0.5.3 adobe_c2pa/0.2.1 c2pa-rs/0.16.0",
      "title": "2304_3Key_AFCEA_BS70_2023_dalibor_CZCERTAINLY.jpg",
      "format": "image/jpeg",
      "instance_id": "xmp:iid:7f39baac-6104-43ff-a614-cbed1765d93a",
      "thumbnail": {
        "format": "image/jpeg",
        "identifier": "adobe-urn-uuid-89969422-4c2b-4ff7-ac09-6b7fc148e7d3.jpg"
      },
      "ingredients": [
        {
          "title": "2304_3Key_AFCEA_BS70_2023.jpg",
          "format": "image/jpeg",
          "document_id": "xmp:iid:f80bde1c-054a-4f76-a71b-80d4583ce541",
          "instance_id": "xmp:iid:f80bde1c-054a-4f76-a71b-80d4583ce541",
          "thumbnail": {
            "format": "image/jpeg",
            "identifier": "xmp-iid-f80bde1c-054a-4f76-a71b-80d4583ce541.jpg"
          },
          "relationship": "parentOf",
          "metadata": {
            "dateTime": "2023-05-26T13:06:53.989Z"
          }
        },
        {
          "title": "czcertainly_white_H_349x83.png",
          "format": "image/png",
          "document_id": "xmp:iid:69e97983-38ce-42ec-a763-8ba0bac2145e",
          "instance_id": "xmp:iid:69e97983-38ce-42ec-a763-8ba0bac2145e",
          "thumbnail": {
            "format": "image/png",
            "identifier": "xmp-iid-69e97983-38ce-42ec-a763-8ba0bac2145e.png"
          },
          "relationship": "componentOf",
          "metadata": {
            "dateTime": "2023-05-26T13:37:22.414Z"
          }
        }
      ],
      "credentials": [
        {
          "@context": [
            "https://www.w3.org/2018/credentials/v1"
          ],
          "credentialSubject": {
            "id": "did:adobe:f41f6b4eb056720f41651ba2a301979568c36a22267b09c490ce79f4a875ba02
9",
            "name": "Roman Cinkais"
          },
          "proof": {
            "created": "2023-05-26T13:06:47.685720796Z",
            "jws": "f298be7a3cc226c030c32a3d847d27af1412d1737adc86819449947f3c207c44b02a034c7"
          }
        }
      ]
    }
  }
}
```

***“Transparentnost a autenticita musí
být základem každé informace a
ekosystému!”***



3Key Company s.r.o.

Roman Cinkais

roman.cinkais@3key.company