



Liccium

KI-Training:
Möglichkeiten des Rechtevorbhalts (Opt-out)
im Rahmen der Text und Data Mining (TDM)
Schranke nach §44b UrhG

Do not train – Nutzung des neuen ISCC zum Ausdruck von Rechtevorbhalten auf Titelbasis

20. Juni 2024 Sebastian Posth
Liccium.com, sebastian@liccium.com

Supported by





SEBASTIAN POSTH

- Unternehmer und Berater zu Innovationsthemen in den digitalen Medienbranchen
- Hintergrund in der Verlagsbranche mit Schwerpunkt auf dem Vertrieb digitaler Medien (Bertelsmann/Arvato, et al.)
- Gründer Liccium.com und Initiator des CreatorCredentials.com Projekts
- Co-Initiator des ISCC
- Gründer and Mitglied des Board of Director's der ISCC Foundation (NL), 2019-2024
- Convenor von ISO/TC 46/SC 9 zum ISCC



Liccium



SEBASTIAN POSTH



- Non-Profit Organisation
- Unterstützt Nutzung und fördert die Weiterentwicklung des ISCC
- ISCC Foundation ist der technische Ansprechpartner für den ISCC



- Gründer und GF von Liccium B.V. (NL)
- [Liccium.com](https://www.liccium.com) and [CreatorCredentials.com](https://www.CreatorCredentials.com) – **Anwendungen des ISCC**, gefördert von der EU Commission
- Liccium unterstützt Rechteinhaber bei der Deklarationen von Rechten und Metadaten (opt-out)



§ 44b TEXT UND DATA MINING

*"Ein **Nutzungsvorbehalt** bei online zugänglichen **Werken** ist nur dann wirksam, wenn er in **maschinenlesbarer Form** erfolgt."*

1. Über welche **Werke** sprechen wir?
2. Wo ist der **Nutzungsvorbehalt** aufzufinden?
3. Wie ist der Nutzungsvorbehalt für **Maschinen interpretierbar**?



THEMEN

1. Identifikation von Mediendateien
2. Verknüpfung von Metadaten
3. Ausdruck von Rechten



IDENTIFIKATION VON MEDIENDATEIEN





IDENTIFIKATION VON MEDIENDATEIEN

Wie identifiziert man Hunderte von Millionen Dateien, die zu Trainingszwecken genutzt werden?

- Standard-Identifizierer (ISBN, ISRC, DOI)?
- Dateinamen?
- Andere Markierungen?

Etablierte Standard-IDs sind **online** nur **selten verfügbar**, werden den Dateien in oftmals **schwer zugänglichen** Datenbanken **manuell** zugewiesen und verwaltet.

1. Like_a_Rolling_Stone.mp3	31. sOnG0987.mp3
2. PopHits_Vol3.mp3	32. 9zXyKlmn.mp3
3. 98KlMnRt.mp3	33. 56GHTyu9.mp3
4. Comfortably_Numb.mp3	34. Born_to_Run.mp3
5. Bohemian_Rhapsody.mp3	35. jKl456Mn.mp3
6. 1wXyZ678.mp3	36. 9jKlMn45.mp3
7. 1234abcd.mp3	37. mUsiC2022.mp3
8. 7MnOpQrt.mp3	38. xYz12345.mp3
9. xYz7838K.mp3	39. RockAnthem_1978.mp3
10. Purple_Haze.mp3	40. aBc123xy.mp3
11. Yesterday.mp3	41. Imagine.mp3
12. JazzFusion_2021.mp3	42. test.mp3
13. Wonderwall.mp3	43. DSM.mp3
14. lmnOp098.mp3	44. 45. Metallica_Live1992.mp3
15. Redemption_Song.mp3	45. Testtest.mp3
16. Sweet_Child_O_Mine.mp3	46. Let_it_Be.mp3
17. mUslc123.mp3	47. Billie_Jean.mp3
18. 6GhYtKlP.mp3	48. 45uloPq2.mp3
19. sOnG4567.mp3	49. What's_Going_On.mp3
20. Enter_Sandman.mp3	50. xYz12345.mp3
21. 34TgVbNm.mp3	51. hat's_Going_On.mp3
22. ClassicBeats_1985.mp3	52. xYz12345.mp3
23. I_Want_to_Break_Free.mp3	53. Green_Grass.mp3
24. mNo45678.mp3	54. Blue_Sky.mp3
25. Smells_Like_Teen_Spirit.mp3	55. Thunder_Road.mp3
26. RockAnthem_1978.mp3	56. Pink_Floyd.mp3
27. Stairway_to_Heaven.mp3	57. Nirvana_Live.mp3
28. uloP6789.mp3	58. Beat_It.mp3
29. 5tGhJkLp.mp3	59. Queen_Concert.mp3
30. Hey_Jude.mp3	60. Firework.mp3




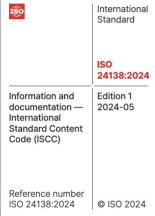
DEZENTRALE IDENTIFIKATION

*"Wie können sich **zwei Parteien**,
die sich **nicht kennen**,
die **keinen Vertrag** abgeschlossen haben,
darüber austauschen, zu **welchem Inhalt**
es einen **Nutzungsvorbehalt** gibt?"*



ISO 24138:2024

 Menu



ISO 24138:2024
Information and documentation
— International Standard
Content Code (ISCC)

Published (Edition 1, 2024)

[Read sample](#)

ISO 24138:2024

Information and documentation
— International Standard
Content Code (ISCC)

Published (Edition 1, 2024)

ISO 24138:2024

CHF **151**

Format
PDF + ePub

Language
 English

[Add to cart](#)

Convert Swiss francs (CHF) to [your currency](#)

Abstract

This document specifies the syntax and structure of the International Standard Content Code (ISCC), as an identification system for digital assets (including encodings of text, images, audio, video or other content across all media sectors). It also describes ISCC metadata and the use of ISCC in conjunction with other schemes, such as DOI, ISAN, ISBN, ISRC, ISSN and ISWC.

General information

Status : Published
Publication date : 2024-05
Stage : International Standard published [60.60]

Edition : 1
Number of pages : 33



Die Internationale Organisation für Normung (ISO) setzt sich aus Vertretern der nationalen Normungsorganisationen zusammen, die Standards für Verlagsbranchen entwickelt, wie z.B. die ISBN, DOI, ISRC, ISWC, ISAN.



ISO 24138:2024

ISCC ist ein offenes System zur dezentralen Identifikation von digitalen Mediendateien ("asset") aller Medien-Typen und Medien-Formate (Text, Bild, Audio, Video).

Das Projekt wurde 2018 von DIN initiiert.

2019 begann die Arbeitsgruppe in ISO/TC 46/SC 9 mit ihrer Arbeit. Der ISCC ist insofern ein Schwesterstandard zu ISBN, DOI, ISRC, etc.

Im Mai 2024 haben Experten aus 29 Ländern aus allen Medienbereichen den ISCC als globalen ISO Standard für die Identifikation digitaler Mediendateien verabschiedet.



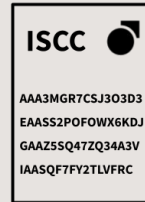
Die Internationale Organisation für Normung (ISO) setzt sich aus Vertretern der nationalen Normungsorganisationen zusammen, die Standards für Verlagsbranchen entwickelt, wie z.B. die ISBN, DOI, ISRC, ISWC, ISAN.



Algorithm



Digital Media Content Asset Codes



AAA3MGR7CSJ303D3
EAASS2POFOWX6KDJ
GAAZ5SQ47ZQ34A3V
IAASQF7FY2TLVFR



THE DNA OF YOUR DIGITAL CONTENT

ISCC:KADV5PDFXBL7HGBXFFW64KVNP6UGTUZC2CJTDBKMFYTTZPLQQVX22FI

Meta-Code

Content-Code

Data-Code

Instance-Code

AAAV5PDFXBL7HGBX

EAASS3P0FKWX7KDJ

GAA5GIWQSMYYKTBO

IAAS0PF50CCW7LIV

**Metadata
Similarity**

**Content
Similarity**

**Data
Similarity**

**Data
Integrity**

Similarity-preserving hashes
(SIM hash)

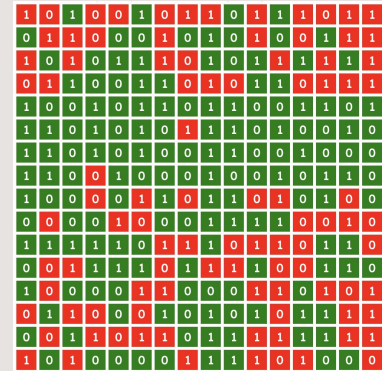
Integrity verifying
checksum (crypto hash)



BEISPIEL: FAKE NEWS

KEC2L03CU6XL6ZVXS3G5LUWUZDEJNBWUBDZPXNR542DDKYVPG372D2A

KECX4FAQRQG2VJ3YS3G5JUWUZDMJNE7GIT67RT27H3AAFVV4NDGABUI



JPG File

Components

JPG File

Similarity

AAA2L03CU6XL6ZVX
EEAZNTOV2LKMRS EW
GAAYNVAI6L53MPPG
IAAYMNL CV43P7IPI

- Meta-Code
- Content-Code
- Data-Code
- Instance-Code

AAAX4FAQRQG2VJ3Y
EEAZNTOU2LKM RWEW
GAAZH ZSE7X4M6XZ6
IAA4AAW WX RUMYAGR

- 41%
- 94%
- 27%
- Unique

Unterschied von 2 bits, bezogen auf den 64 bit hash

Example

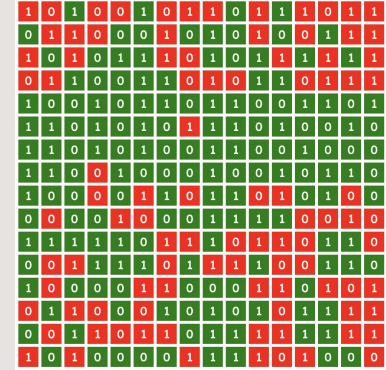


BEISPIEL: FAKE NEWS

KEC2L03CU6XL6ZVXS3G5LUWUZDEJNBWUBDZPXNR542DDKYVPG372D2A

KECX4FAQRQG2VJ3YS3G5JUWUZDMJNE7GIT67RT27H3AAFVV4NDGABUI

Das Clustering und Matching ähnlicher Inhalte ist auch dann möglich, wenn man nur den Zugriff auf die ISCC Codes hat!



JPG File

Components

JPG File

Similarity

AAA2L03CU6XL6ZVX
EEAZNTOV2LKMRS EW
GAAYNVAI6L53MPPG
IAAYMNL CV43P7IPI

Meta-Code
Content-Code
Data-Code
Instance-Code

AAAX4FAQRQG2VJ3Y
EEAZNTOU2LKM RWEW
GAAZH ZSE7X4M6XZ6
IAA4AAW WX RUMYAGR

41%
94%
27%
Unique

Unterschied von 2 bits, bezogen auf den 64 bit hash

Example



UNTERSTÜTZTE MEDIEN/DATEI-FORMATE

- **TEXT**

pdf, doc, docx, xls, xlsx, pptx, epub, html, xhtml, odt, pdf, rtf, txt, xml, json, md

- **IMAGE**

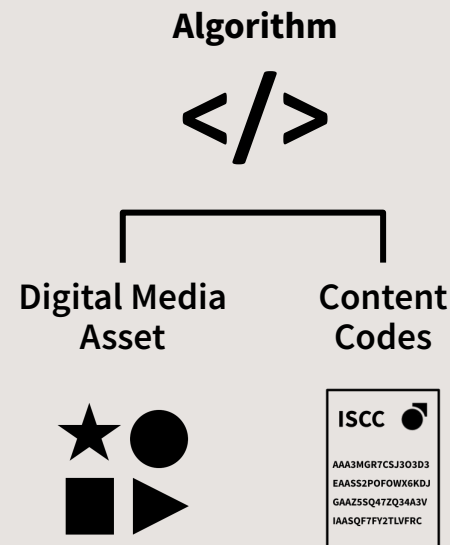
gif, jpg, png, tif, bmp, psd, eps, webp, avif, heic

- **AUDIO**

aif, flac, mp3, opus, ogg, wav

- **VIDEO**

3gp, 3g2, asf, avi, drc, flv, f4v, flu, gif, h264, mpg, mp4, mkv, mov, ogv, rm, swf, webm, wmv



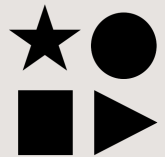


Algorithm



**Digital
Media Asset**

**Content
Codes**



ISCC 
AAA3MGR7CSJ303D3
EAASS2POFOWX6KDJ
GAAZ5SQ47ZQ34A3V
IAASQF7FY2TLVFRC

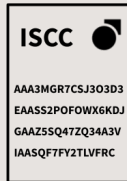
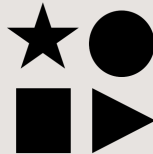


Algorithm



Digital
Media Asset

Content
Codes



Opt-out





VERKNÜPFUNG VON METADATEN

Metadata Viewer: [Flowers (2)] (imported)

Exif XMP IPTC

IPTC Tag	Value
Iptc.Application2.Byline	Sebastian Posth
Iptc.Application2.Copyright	(c) 2023 Sebastian Posth
Iptc.Application2.Credit	All rights reserved
Iptc.Application2.Keywords	Field; Flowers, Randstad
Iptc.Application2.RecordVersion	4

Help Close



VERKNÜPFUNG VON METADATEN

WATERMARKING / STEGANOGRAPHY

Daten werden sichtbar/
unsichtbar eingebettet in
die nicht/sichtbaren Ebenen
der Mediendatei

LOCATION / DOMAIN-BASED

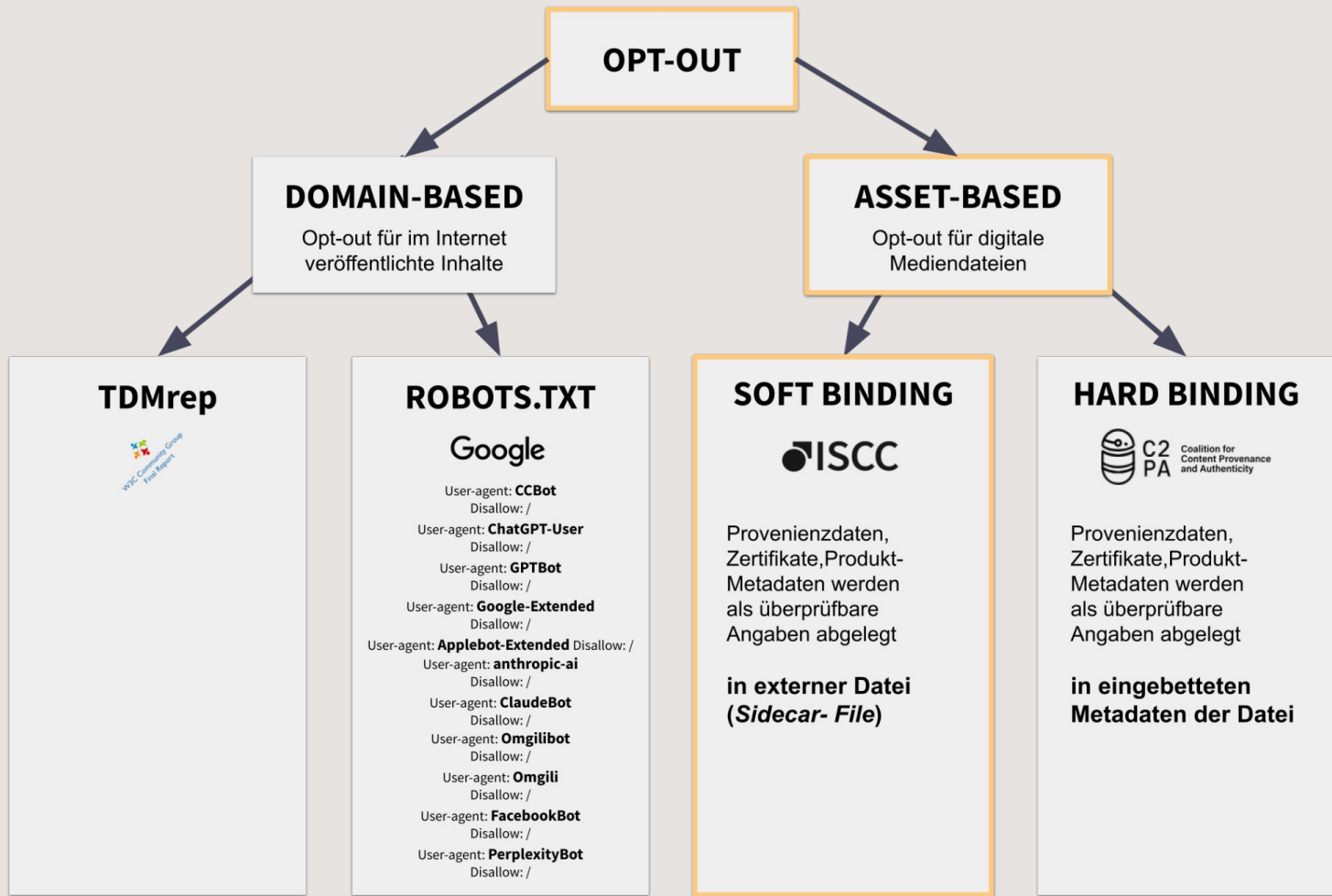
Metadaten werden in
robots.txt-Datei untergebracht
oder in HTML/HTTP Metadaten
der Domain

METADATA EMBEDDING / HARD BINDING

Metadaten werden unmittelbar
eingebettet in die Mediendatei

EXTERNAL / SOFT BINDING

Metadaten werden in externer
Datei bereitgestellt und verlinkt
von ISCC

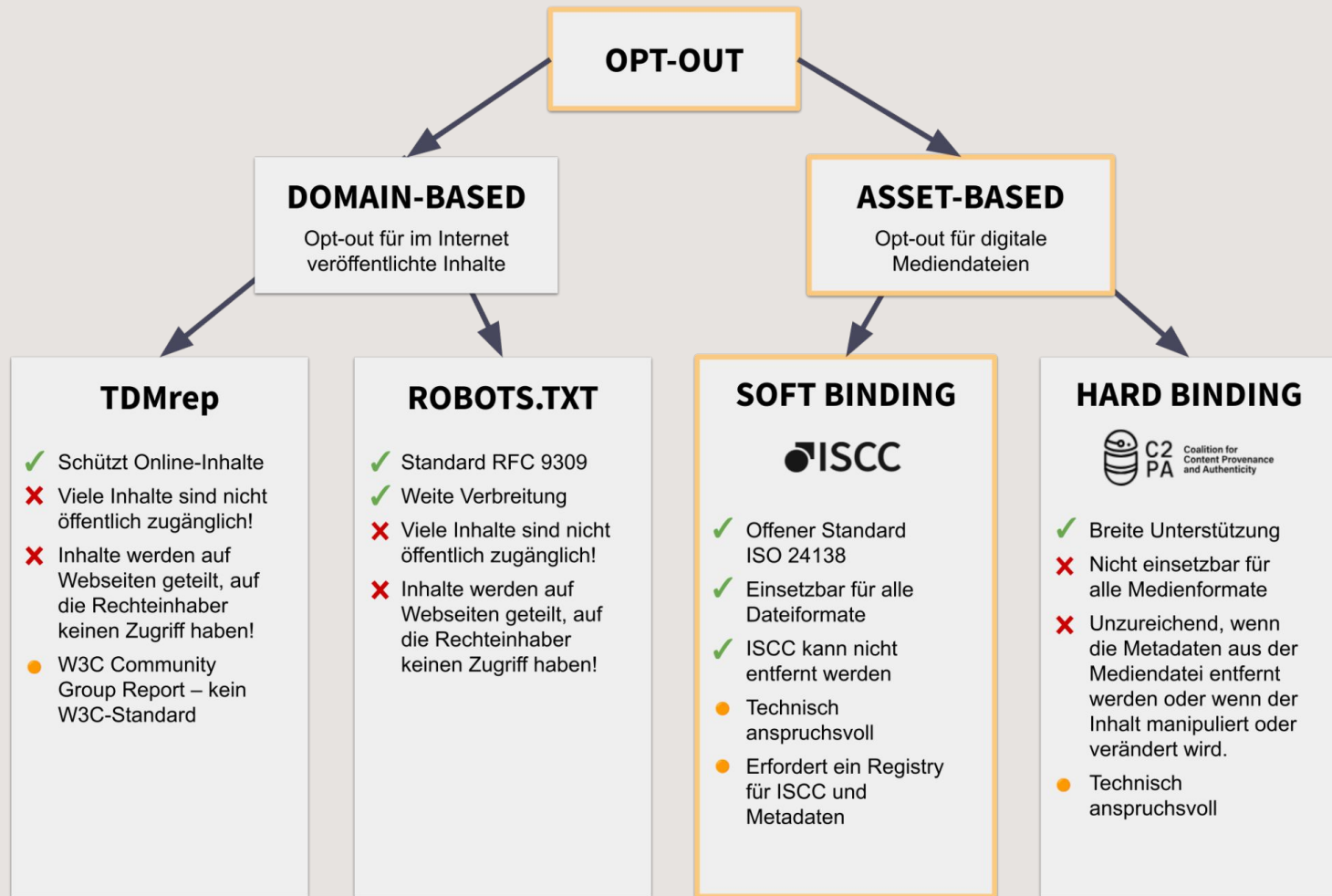




PERSISTENTE VERKNÜPFUNG VON METADATEN

Warum brauchen wir **Soft-Binding**?

1. Der Inhalt wird verändert oder manipuliert
2. Eingebettete Metadaten werden aus der Mediendatei entfernt
3. Wasserzeichen oder steganografische Daten werden entfernt
4. Der Inhalt wird auf Webseiten geteilt, über die die ursprünglichen Rechteinhaber:innen keine Kontrolle haben
5. Der Inhalt wird in den soziale Medien geteilt





EMPFEHLUNG FÜR OPT-OUT-ERKLÄRUNGEN

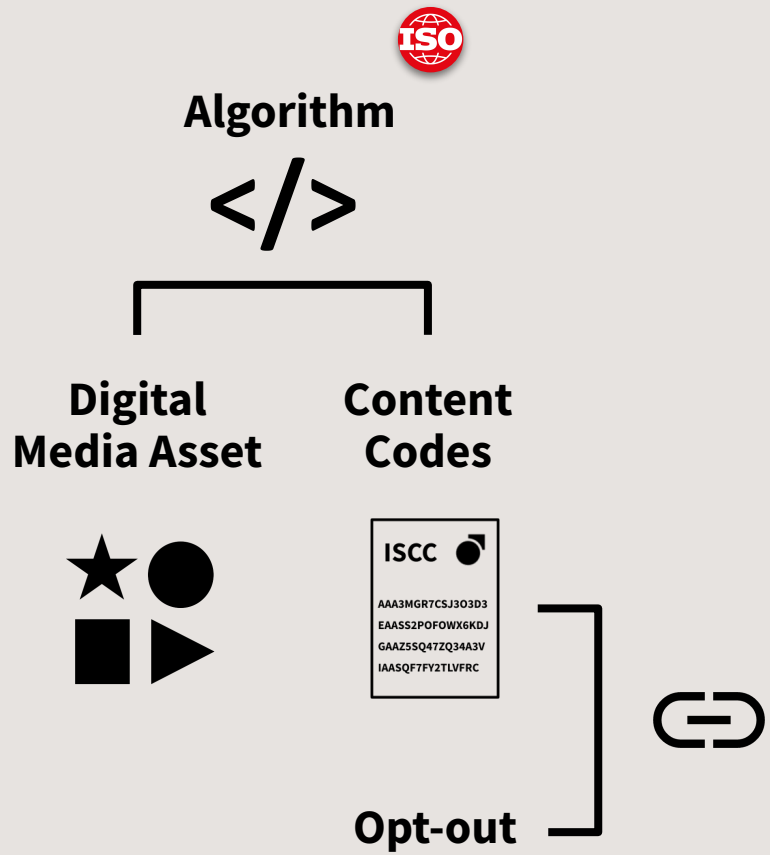
- Maschinenlesbar (Rechtliche Anforderung)
- Untrennbar mit dem Inhalt ("Asset") verbunden, robust bei Social Sharing und digitale Verbreitung
- Resilient gegenüber der Manipulation von Inhalten
- Zuverlässig bei Entfernung eingebetteter Metadaten, Wasserzeichen und steganografischer Daten
- Nutzung digitaler Signaturen und Zertifikate (Attribution)
- Nutzung verifizierbarer Zeitstempel
- Basierend auf internationalen Standards (ISO, W3C)



AI OPT-OUT

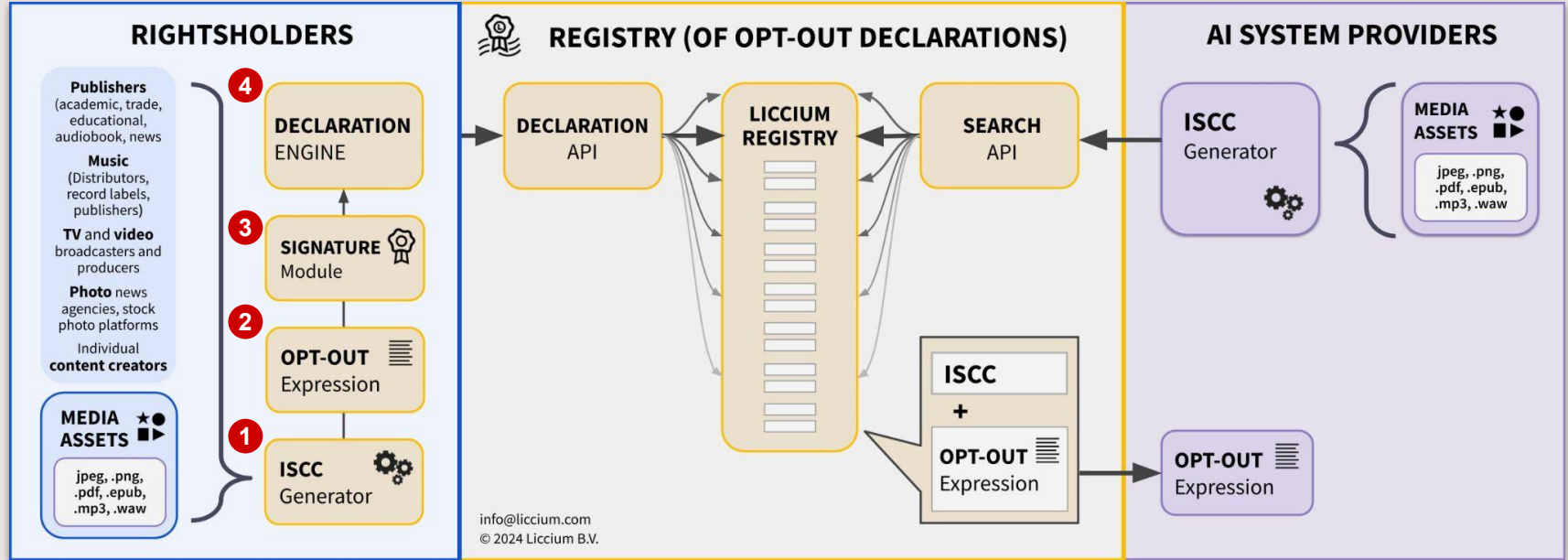
Urheber:innen und Rechteinhaber:innen können eine maschinenlesbare Opt-out-Erklärung untrennbar mit dem ISCC verküpfen, um zu verhindern, dass ihre Inhalte zu KI-Trainingszwecken verwendet werden.

Anbieter von KI-Anwendungen können die **Opt-out Erklärungen aus dem ISCC herleiten** und so die Vorbehalte der Rechteinhaber respektieren.





AI OPT-OUT USE CASE

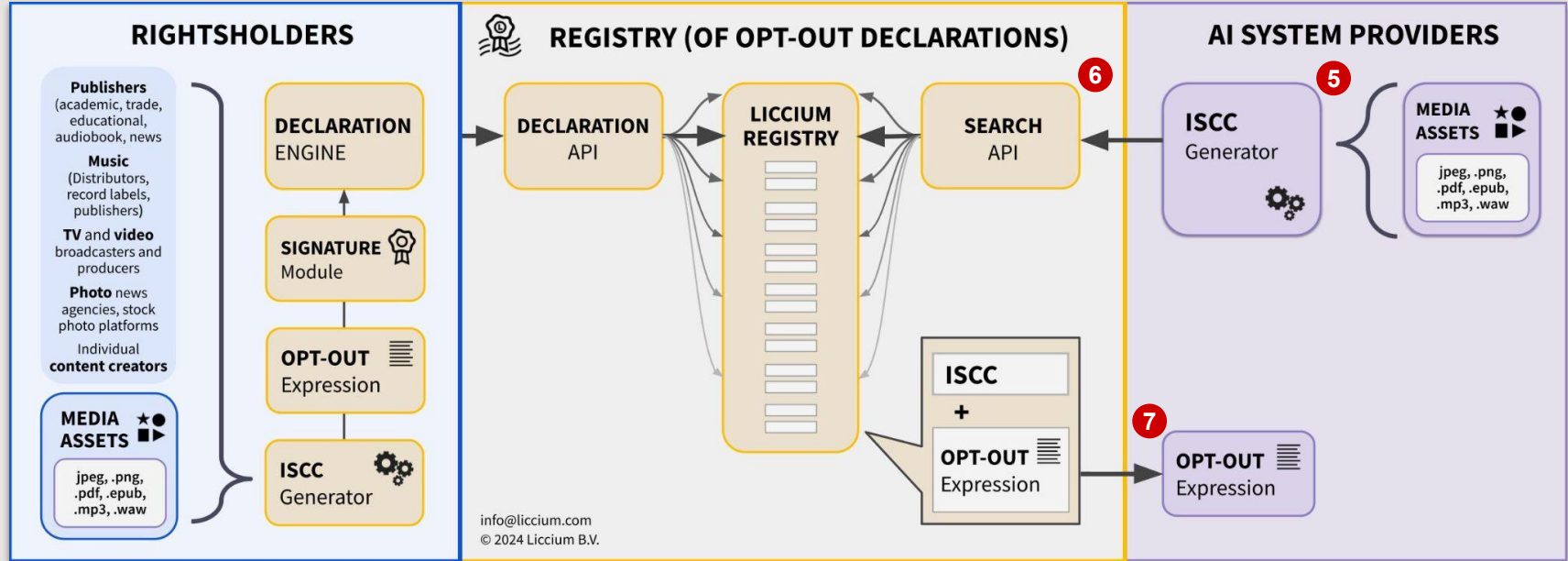


- 1 ISCC wird aus Datei generiert
- 2 Opt-out Kennung wird an ISCC gebunden

- 3 Deklaration wird digital signiert
- 4 Deklaration wird zum Registry geschickt



AI OPT-OUT USE CASE



- 5 ISCC wird aus Datei generiert
- 6 AI Anbieter sucht ISCC in Registry

- 7 AI Anbieter findet Opt-Out Erklärung für gleiche oder ähnliche Inhalte



ZUSAMMENFASSUNG VORTEILE TDM·AI / ISCC

STÄRKEN

- Der ISCC ist ISO-Standard
- Der ISCC kann von jeder Nutzer:in mit Open-Source-Software aus der Datei erzeugt werden
- Der ISCC kann nicht aus dem Inhalt entfernt werden
- Der ISCC kann gleiche oder ähnliche Inhalte auffinden
- Kompatibel mit allen anderen Ansätzen (C2PA, IPTC, etc.)

VORAUSSETZUNGEN

- 'Soft-Binding' bedingt ein Registry für ISCC and Rechte/Metadaten
- ISCC must generiert werden, Rechte/Metadaten müssen verknüpft werden und es muss eine öffentliche und digital unterschriebene Deklaration geben
- Das Auffinden ähnlicher Dateien ist technisch anspruchsvoll und bedarf der Interpretation



AUSDRUCK VON RECHTEN





TDM·AI

TDM·AI OPT-OUT PROTOKOLL

TDM·AI ist ein Protokoll zur Verknüpfung maschinen-lesbarer Opt-out-Erklärungen für KI-Trainingsdaten mit digitalen Mediendateien.

Es basiert auf der DSM-Directive unter Nutzung der Vorteile des International Standard Content Code (ISCC) und Creator Credentials.

Gemeinsam entwickelt von **Sabine Richly**
(LL.M.- & MBA) <https://medialaw.digital/>



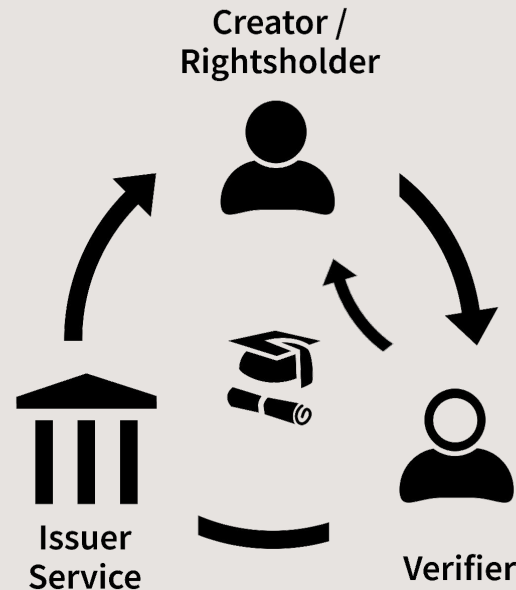


CREATOR CREDENTIALS



Verifizierbare Nachweise (VCs) dienen als Mittel zur Identifizierung von Urhebern:innen und Rechteinhaber:innen.

Liccium erleichtert ihnen die Beantragung, Verwaltung und Überprüfung von Zertifikaten und verifizierbaren Nachweisen (VCs) und stärkt dadurch Vertrauen in Deklarationen und Attribution.





AUSDRUCK VON RECHTEN (TDM·AI)

Opt-out Declaration

```
1  {
2    "TDMAI": false,
3    "TDMAI_summary": "Content must not be used for AI training purposes.",
4    "TDMAI_policy": "Automated analysis of the work to extract information from it,
                    especially about patterns, trends, and correlations for the purpose of training
                    models and applications of generative AI, is reserved. Text and Data Mining (TDM) is
                    permitted for general purpose AI systems that do not generate synthetic audio,
                    image, video, or text content and for scientific research purposes or for temporary
                    acts of reproduction as provided for in Article 5(1) of Directive 2001/29/EC."
5  }
```



METADATEN DER DEKLARATION



```
{
  "@context": "http://purl.org/iscc/context/0.4.0.jsonld",
  "$schema": "http://purl.org/iscc/schema/0.4.0.json",
  "iscc": "KECUV7T2M0Y4UDAX53VVK4KFBIF7J5LOWG0U3QYGWU27JN0XEB2S5A",
  "original": true,
  "version": 52,
  "entryUUID": "d3132443-40c0-48e1-96f5-7442a8bec7af",
  "credentials": "https://www.liccium.dev/api/v1/credentials/e3acd695-5008-48b1-9109-1828dd5cc626",
  "liccium_plugins": {
    "tdmai": {
      "TDMAI": false,
      "TDMAI_policy_URL": "https://olive-labour-unicorn-913.mypinata.cloud/ipfs/QmX8qUyi2iXRT7BEX6D2SSJUqngQBN1hWkA1TFbQ9Zisad"
    },
    "iptc": {
      "digitalsourcetype": "minorHumanEdits",
      "keywords": "Field; Flowers, Randstad, NGI, Trustchain, 11:48",
      "creator": "Sebastian Posth",
      "credit": "All rights reserved",
      "creditText": "All rights reserved",
      "copyrightNotice": "(c) 2023 Sebastian Posth"
    },
    "c2pa": {
      "active_manifest": {
        "label": "b36679d8-99de-4610-bc1a-78aea6c10bb8",
        "assertions": [
          {
            "data": {
              "entries": {
                "cawg.ai_training": {
                  "use": "notAllowed"
                },
                "cawg.data_mining": {
                  "use": "notAllowed"
                },
                "cawg.ai_inference": {
                  "use": "notAllowed"
                },
                "cawg.ai_generative_training": {
                  "use": "notAllowed"
                }
              }
            }
          }
        ]
      }
    }
  }
}
```




AUSDRUCK VON RECHTEN (C2PA)

```
{
  "entries":
    "cawg.ai_training": {
      "use": "allowed"
    },
    "cawg.ai_generative_training": {
      "use": "notAllowed"
    },
    "cawg.data_mining": {
      "use": "constrained",
      "constraint_info": "may only be mined on days whose names end in 'y'"
    }
}
```



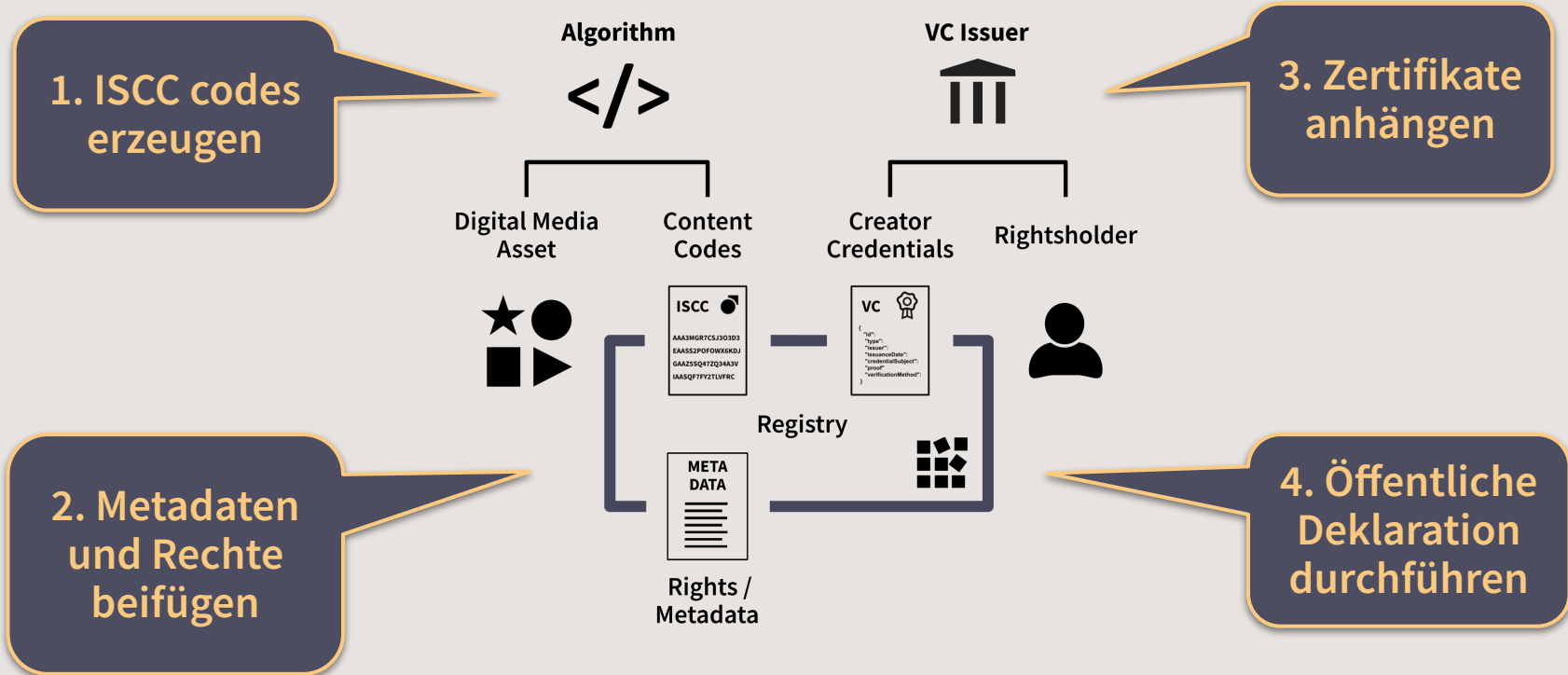
AUSDRUCK VON RECHTEN (TDMrep)

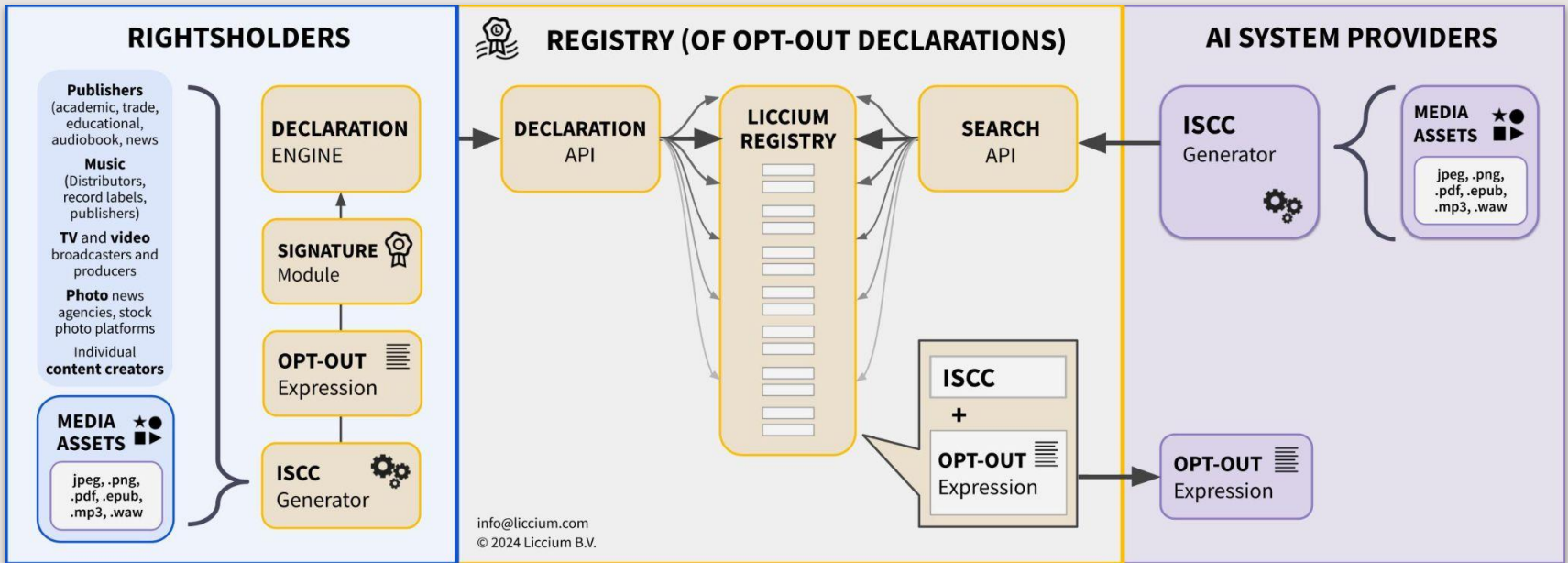
EXAMPLE 1

```
[  
  {  
    "location": "/",  
    "tdm-reservation": 1  
  }  
]
```



HOW CREATORS REGAIN CONTROL







Liccium

KI-Training:
Möglichkeiten des Rechteevorbehalts (Opt-out)
im Rahmen der Text und Data Mining (TDM)
Schranke nach §44b UrhG

Vielen Dank!

Liccium.com

sebastian@liccium.com

Supported by

