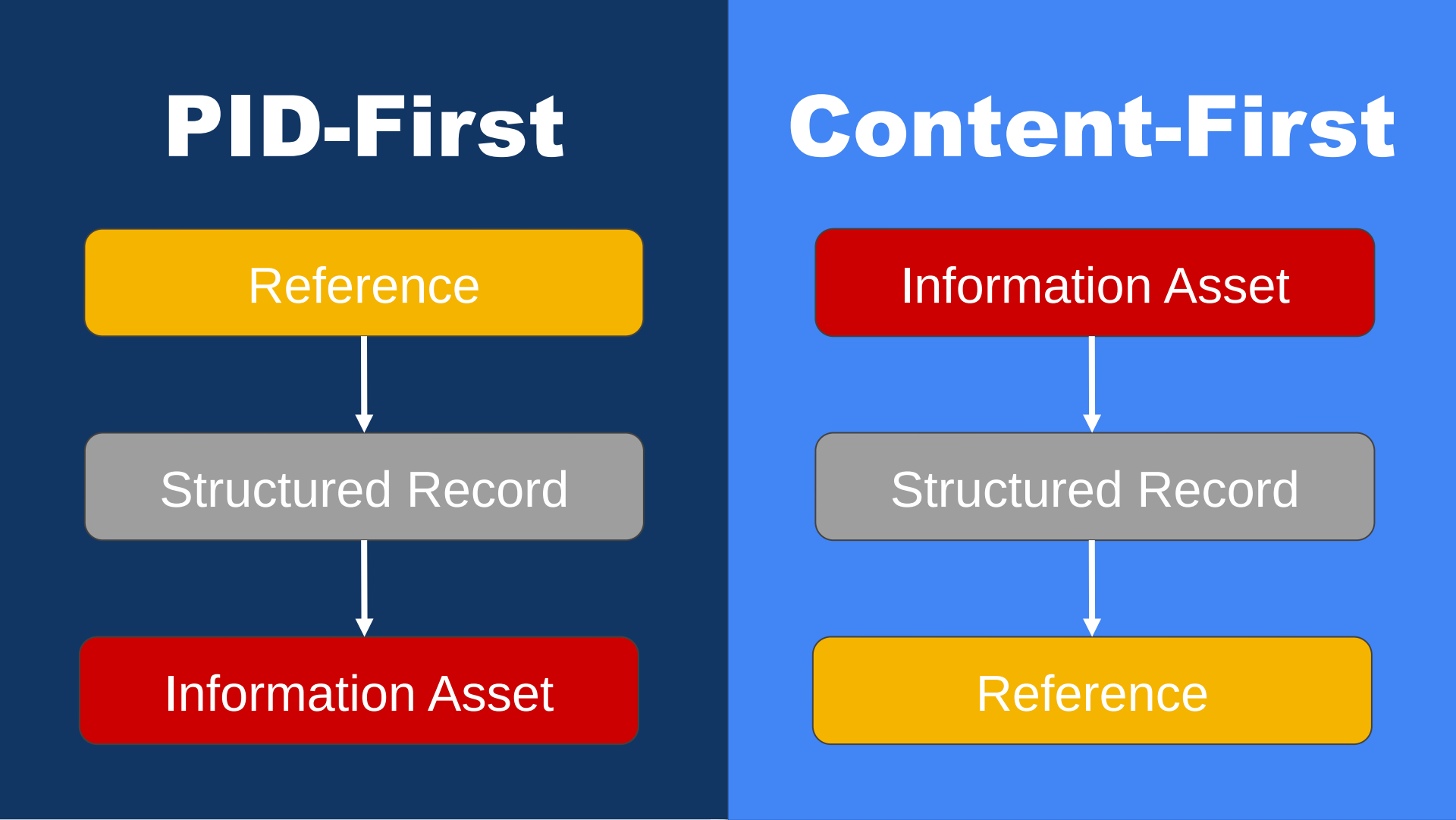
ISCC

From Content Fingerprint to FAIR Digital Object

The ISCC Protocol as a Content-Based Discovery Layer for the FDO Ecosystem

Titusz Pan
Dr. Martin Etzrodt

ISO 24138:2024 · ISCC Foundation · FDO Forum 2026 · TU Wien, Vienna


PID-First
Reference
Structured Record
Information Asset
Content-First
Information Asset
Structured Record
Reference

TC 46 - Information and Documentation
SC 09 - Identification and Description

ISCC

ISO 24138:2024

Conceived
2016-06-29

WG 18 Started
2019-10-29

Published
2024-05-15

ISO 2108:2017	ISBN	International Standard Book Number
ISO 3297:2022	ISSN	International Standard Serial Number
ISO 3901:2019	ISRC	International Standard Recording Code
ISO 15706-1:2023	ISAN	International Standard Audiovisual Number
ISO 15707:2022	ISWC	International Standard Musical Work Code
ISO 27729:2012	ISNI	International Standard Name Identifier
ISO 26324:2022	DOI	Digital Object Identifier System
ISO 24138:2024	ISCC	International Standard Content Code

ISCC

The DNA of your digital content

Estimate similarity using ISCC-CODEs

ISCC: KED572P4AOF5K6QXQA4T6OJD5UGX7UBPFW2TVQNTHBCKFRFCANCZARQ4K6NSFZQSH4GQ

Meta-Code	Semantic-Code	Content-Code	Data-Code	Instance-Code
AAA572P4AOF5K6QX	CEAYAOJ7HER62DL7	EEA5ALZNWU5MDMZY	GAAUJIWEUIBULECG	IAARYV43ELTBEPYN

Abstract & Persistent — Concrete & Volatile

Metadata Similarity	Semantic Similarity	Syntactic Similarity	Data Similarity	Data Integrity

Components are self-describing and can be used standalone or in combination and at different length

Content-Code (Image)

CCDFPFc87Mhd
T

CTWAGYJ9HZGj
1

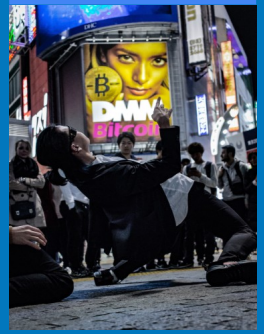
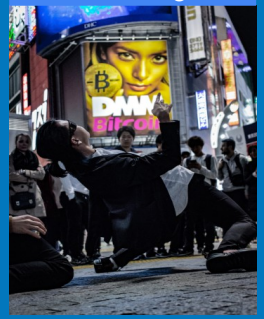
CDhydSjQXDXV
k

CRd5bk4SrBpzt

If we want to identify “Content” we cannot rely on “Data”:

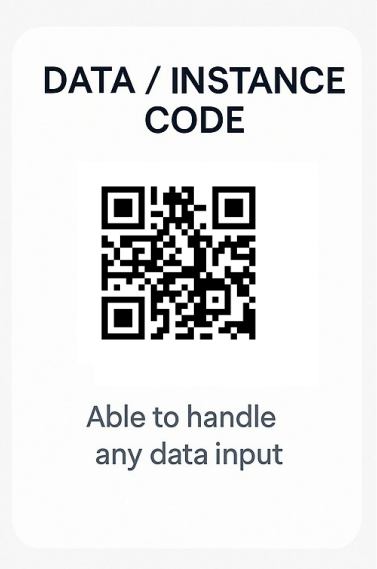
- Two “identical” images
- Yet the data is completely different
- Due to different file formats
- ISCC Content-Code identifies information structure - not raw data

Perceptual Identification

JPG Image	JPG Data	JPG SHA1	JPG Content-ID
	49 74 27 73 20 6e 6f 74 20 61 62 6f 75 74 20 62 61 6e 6b 69 6e 67 20 74 68 65 20 75 6e 62 61 6e 6b 65 64 2e	7b 24 1f 77 f0 f2 96 df 73 b5 e0 38 97 6a 5e 3b d0 12 bd 23	CYHa5UMqq1iQS
=	≠	≠	=
PNG Image	PNG Data	PNG SHA1	PNG Content-ID
	54 68 65 20 43 75 72 72 65 6e 63 79 20 75 73 65 64 20 6f 6e 20 43 6f 62 6c 6f 20 69 73 20 43 68 61 72 6d 2e	7e bd c5 c5 c0 30 d5 4c 30 c0 31 df 4c 9e ff d5 b2 ad e8 2d	CYHa5UMqq1iQS

BIOCODES: ISCC-SUM

High-Performance Implementation for BIO-CODES

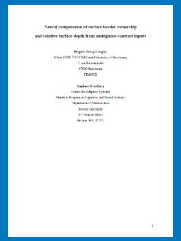


- Optimized ISCC for Scientific TB-scale datasets ~1GB/s
- Implements Data & Instance code
- TREEWALK : deterministic directory hashing for large dataset collections
- READY: Rust library, Python bindings, CLI tool
- Standard compliant checksum

Funded by: OSCARS Open Science Clusters' Action for Research & Society

Collaboration with: LEIDEN CELL OBSERVATORY

Example one **DOI** multiple matching **ISCC**

Paper: Neural Computation of Surface Border Ownership and Relative Surface Depth from Ambiguous Contrast Inputs

Host	DOI	ISCC
hal.archives-ouvertes.fr	10.3389/fpsyg.2016.01102	ISCC:KACXDXP4BDIPANIC2KOV6G5QSBYE6MMBKWG4BIUNAUDBK53OF5IXLQI
arxiv.org	10.3389/fpsyg.2016.01102	ISCC:KACYVXLR2SN2UQNKSKMQWKPI2CZM245MEWE6D3UZNL4DTMKRHFHKP4Q
hal.archives-ouvertes.fr	10.3389/fpsyg.2016.01102	ISCC:KACXDXP4BDIPANIC2KOV6G5QSBYE6MMBKWG4BIUNAUDBK53OF5IXLQI
frontiersin.org	10.3389/fpsyg.2016.01102	ISCC:KACXDXP4BAIS5YEW62KOV6G5QSBYE6MMBKSG4BIUNAUZD43KLLDPZ4UQ

ISCCs can be compared to estimate content similarity

What does ISO 24138:2024 standardize?

- Syntax and structure of the ISCC
- Set of algorithms for various ISCC-UNITs (text, image, audio, video, data, metadata)

What does ISO 24138:2024
NOT standardize (yet)?

- **Who** created an ISCC
- **When** was an ISCC created
- **Where** do I find metadata

ISCC Discovery Protocol

Signing, Timestamping, Discovery

IsccNote Message


```
{
  "iscc_code": "ISCC:KACWN77F73NA44D6EUG3S3QNJIL2BPPQFMW6ZX6C…",
  "datahash": "1e205ca7815a69c1d530176d549b5d18d038eb5280b4b3…",
  "nonce": "000faa3f18c7b9407a48536a9b00c4cb",
  "timestamp": "2025-01-15T12:00:00.000Z",
  "gateway": "https://example.com/iscc-id/{iscc_id}/",
  "units": [
    "ISCC:AADWN77F73NA44D6X3N4VEUAPOW5HJKGK5JKLNGLNFPOESXWYDVDVUQ",
    "ISCC:EADSKDNZNYGUUF5AMFEJLZ5P66CP5YKCOA3X7F36RWE4CIRCBTUWXYY",
    "ISCC:GAD334BLFXWN7QWLCSBGJMLRZW73FFNV7ORVUKN23UWPKGQCWTIHQKY"
  ],
  "signature": {
    "version": "ISCC-SIG v1.0",
    "controller": "did:web:example.com",
    "pubkey": "z6MkmeDbeC5BecFmVnTHA5PWEBaVUrGLdB3weGE2KYnXfHso",
    "proof": "z5j9nrpPw3oYSAN4XbCvk2sUtkwrueTD6V2Y35gS1KFTo…"
  }
}
```


Cryptographic Public Key

Owner's key ensures secure and verifiable ownership.

Timestamp

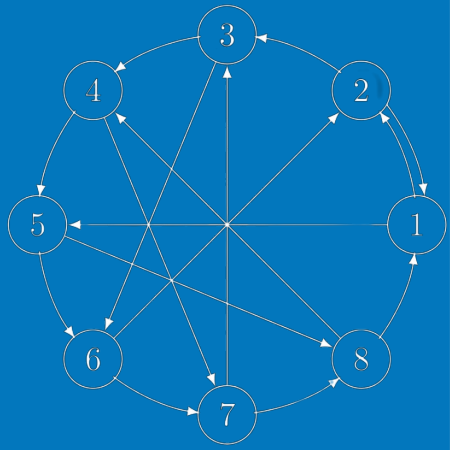
Timestamp provides proof of existence and creation time.

URL Location

URL links to metadata and associated services.

Digital Content

Content represented by ISCC-CODE for identification.



Scalable content based discovery network

The Declaration Event Is the FDO

1. **Generate Ed25519 keypair**
   Self-sovereign, did:web
2. **Generate ISCC-CODE**
   From the digital content
3. **Sign an IsccNote**
   ISCC-CODE + datahash + nonce
4. **Submit to ISCC-HUB**
   Chosen from the federation
5. **Receive ISCC-ID**
   52-bit timestamp + 12-bit HUB-ID
6. **Get IsccReceipt**
   W3C Verifiable Credential

IDP → FDO Mapping		
`Declaration Event`	→	**The FDO**
`ISCC-ID`	→	**The PID**
`IsccReceipt`	→	**The FDO-Record**
`ISCC-CODE`	→	**Metadata attribute**
`Ed25519 public key`	→	**Actor identity**

Content-Based Discovery

From content to FDOs - a new discovery vector for the ecosystem

Three Discovery Modes

ISCC-ID Lookup

Direct PID resolution

ISCC-CODE Discovery

Find similar content via fingerprints

Hash Lookup

Exact binary content match

ISCC

Thank you
for your attention

ISCC – Discovery
Protocol concept

Contact Information:

Titusz Pan
tp@iscc.io

Dr. Martin Etzrodt
me@iscc.io