

# ISCC Test

## *with Images*

Version open for feedback

2020-06-16

Tests performed by Sebastian Posth

## Contact

ISCC Foundation

<https://iscc.foundation>

Boerhaavelaan 99

2334 EG Leiden

Netherlands

Sebastian Posth

Email: [posth@iscc.foundation](mailto:posth@iscc.foundation)

+31 6 12676726

Titusz Pan

Email: [tp@iscc.foundation](mailto:tp@iscc.foundation)

+49 761 2176798

# ISCC – Executive Summary

The ISCC is a new universal identifier for digital content. It helps creative individuals and media organisations to better manage digital content by using state of the art machine learning, fingerprinting and cryptographic technologies. ISCC codes can be used to identify digital assets of all media-types like text, image, audio and video in all granularities across all industry sectors alongside existing standard identifiers.

With modern built-in features like versioning, deduplication, content clustering, timestamping, similarity estimation, granular identifier assignment, and cryptographic proofs it enables transactions to operate faster and more efficiently in networked and increasingly decentralized media environments.

ISCC identifiers can be generated by anyone free of charge and independent of centralised third party services by using open-source software. They are optimized to be registered on blockchains but also used in off-chain environments.

## ISCC Key Features and Differentiators

- Universal identifier for digital text, image, audio, video
- Decentralized identifier issuance by an algorithm
- Bi-directional inherent linking of media-asset and identifier
- Lightweight, multi-faceted fingerprint designed for digital content
- Computational verification of content integrity
- Algorithmic similarity estimation and deduplication

## ISCC Architecture

The ISCC is a content code which is created from the content file itself. Processing the content with the algorithms defined by [ISCC specification](#) creates a unique composite code, which consists of four major elements: the Meta-ID, Content-ID, Data-ID, and Instance-ID codes. Thus, the ISCC identifies and describes content across multiple, hierarchical layers: From the embedded or external metadata, the normalized content, the encoded file data down to the bitstream of an individual file. It can be used to automatically distinguish different versions of the same content, to ensure data integrity, to de-duplicate, or to disambiguate content in a given content repository.

## Example of an ISCC

Meta-ID	Content-ID	Data-ID	Instance-ID
CC7na3ep5Vv4p	CYLTkHXrsxJ62	CDdAbVBuuiQD1	CRbVSQFr8WJut

*Example of an ISCC. See the ISCC specification: <https://iscc.codes>*

# Use Cases for Images

The ISCC for images can be used for a number of use cases:

## Decentralized Content Identification

Anyone with access to digital content can independently generate an ISCC from the content file itself and identify it unambiguously without the need to consult or rely on a centralised authority like a registration agency or service to properly identify the content. The structure of the ISCC supports search and discoverability of content on the web and in distributed network environments.

## Efficient Data Exchange

During content creation, review processes, content collaboration or distribution, same or similar files are being exchanged among various parties. Digital content creators and market participants are sending and receiving digital content files from multiple sources to multiple recipients.

When digital content has been ingested by a receiving party it is not always possible:

- To identify digital content files and distinguish between newly received content files and content files that have been received at an earlier point in time,
- To tell whether the exact same digital content file (duplicate content) has been ingested already, or
- Whether the digital content file has been changed (modified or manipulated version), or
- Whether the file name correctly describes the content or content update.

The ISCC supports efficient data exchange and interoperability through algorithmic similarity estimation and deduplication.

## Data Deduplication and Disambiguation

Professional photographers, archives or other repositories are dealing with a large number of images in various versions, editions, resolutions, file formats, and sizes (e.g. for print and web). Same or similar photos are also assembled in various content collections for various users, customers or other purposes. From all images, multiple backups are being produced and redundantly stored on hardware devices or cloud services. The hosting of duplicate images requires a lot of hardware resources and bandwidth.

Apart from the use of commercial image recognition software, complex file naming patterns are a common way to organize the content repositories. But proprietary file naming (Example: *[Date] [Original Camera File Name] [Name/Topic] [Edition] [Resolution] [File Format]*) is time-consuming and prone to errors. .

The ISCC supports automated content clustering and de-duplication.

## Integrity Verification and Content Versioning

With the ISCC registered on any public ledger it is possible to timestamp content versions and variants in order to claim rights or to create an auditable history of related documents over time. Timestamping allows to verify content integrity and recognize changes or modifications. It helps to identify content versions on a time scale in order to make sure that users are referring to the correct same file or related versions of the same content.

## Purpose of the Test

It is the purpose of this test to provide examples of the ISCC for images, compare two ISCC's of similar content files to demonstrate near duplicate content matching capabilities (lightweight fingerprinting) and the robustness of the ISCC against changes in format, compression, color modifications and image manipulations like crops or erased content elements.

Furthermore, we greatly appreciate any feedback to this document. Please feel free to reach out with questions or comments.

# Testing Resources

## ISCC Specification Version 1.x

The test is based on the [specification in version 1.x](#). This is the latest in-development version of the ISCC Specification. While there is already a Version 1.0 specification, we are still expecting backward incompatible changes until Version 2.0. We encourage partners to follow development and test, implement, and give feedback based on this latest version of the ISCC Specification.

## Testing Tools

### ISCC Online Demo

An online demo where users can generate ISCC codes for digital media files:

<https://iscc.coblo.net/>

### Command Line Tool – iscc-cli

A command line tool that generates ISCC Codes for digital media files:

<https://github.com/iscc/iscc-cli>

### Jaccard-Similarity

The similarity of the Meta-ID, Content-ID and Data-ID components is measured in percent. The jaccard coefficient measures similarity of sets. It is calculated by dividing the number of equal elements (intersection) by the number of total unique elements (set union). The similarity of ISCC components is an estimate of the similarity of the metadata, content or data. The visualization in the online demo shows the bit-by-bit difference between ISCC codes. Each bit that differs in comparison to the previously generated ISCC code is marked red.

## Data Sources

### Images Referencing

For the tests, two images have been used:

<https://pixabay.com/nl/illustrations/natuur-zomer-lake-reflectie-boom-4400180/>

<https://pixabay.com/de/photos/b%C3%A4ume-h%C3%BCgel-gr%C3%BCn-blau-natur-park-790220/>

To explain how the Content-ID for images works (see below), the following image has been used:

<https://morguefile.com/p/121925>

### Test Images

All versions and variants can be downloaded, here:

[https://drive.google.com/drive/folders/12dUY1DpQ13Dijfxp7\\_DcV4GNM-SrtpXx?usp=sharing](https://drive.google.com/drive/folders/12dUY1DpQ13Dijfxp7_DcV4GNM-SrtpXx?usp=sharing)

### Software for Conversions and Image Manipulations

For the conversions, image modifications and manipulations, the following software tools have been used:

GIMP (for Mac and Windows) and Preview (for Mac)

### Remark

The following tests were created manually from the perspective of a business user and are not to be understood as a comprehensive scientific or technical performance benchmark.

## The Meta-ID for Images

The Meta-ID for images is generated from the embedded title or external basic metadata. In the following tests, no title has been provided and embedded in the images. In these cases, the Meta-ID is generated from the normalized file name minus file extension.

# The Content-ID for Images

How does the Content-ID for images, the second component of the ISCC work?

While the [Meta-ID is generated from the embedded or external basic metadata of the content](#), the [Instance-ID](#) and the [Data-ID](#) are created from the raw data of the media file, the [Content-ID for images](#) is generated from extracted and normalized grayscale pixel data.

With regard to a generic deployability of the ISCC the specification suggests an algorithm to generate an ISCC for images that has moderate computation requirements and is easy to implement while still being robust against common image manipulations: the perceptual hash algorithm. For a technical deep dive and a comparison of the various hashing algorithms for images, see the blog post [Testing Different Image Hash Functions](#), from which the following examples are taken.

The perceptual hash algorithm initially takes a colour image, calculates a gray value image from it and scales it down to a  $32 \times 32$  image. To this down-scaled image a discrete cosine transform (DCT) is applied, first per row and afterwards per column.

$$\text{Discrete cosine transform (DCT): } X_k = \sum_{n=0}^{N-1} 2n * \cos(\Pi * k * \frac{2n+1}{2N}) \quad \forall k \in [0, N]$$

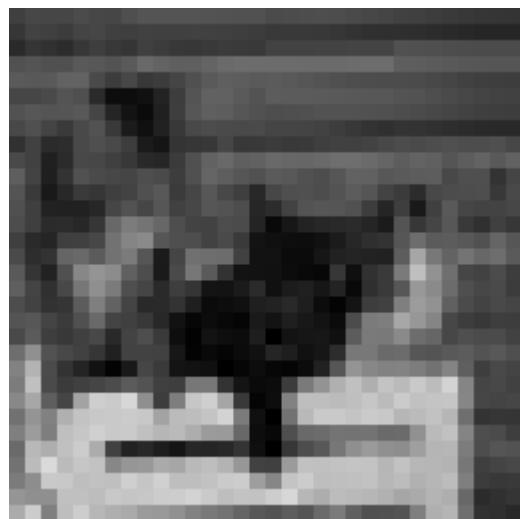
The pixels with high frequencies are now located in the upper left corner. Now the image is cropped to the upper left  $8 \times 8$  pixels. From these 64 pixels, the median of the gray values is calculated. Analogous to the median hash algorithm, a hash value from the image is generated.



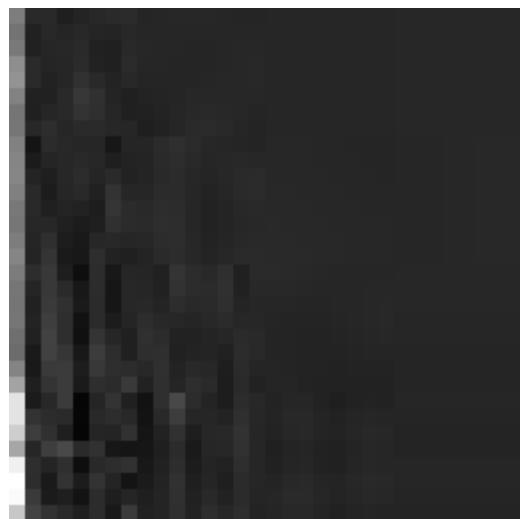
Source image



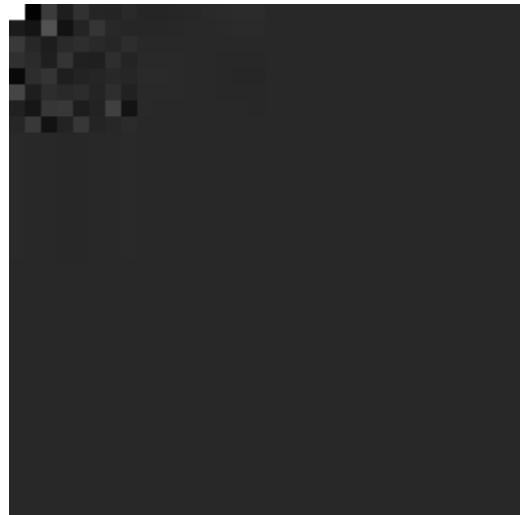
Grayscale image



Gray value image, size 32×32



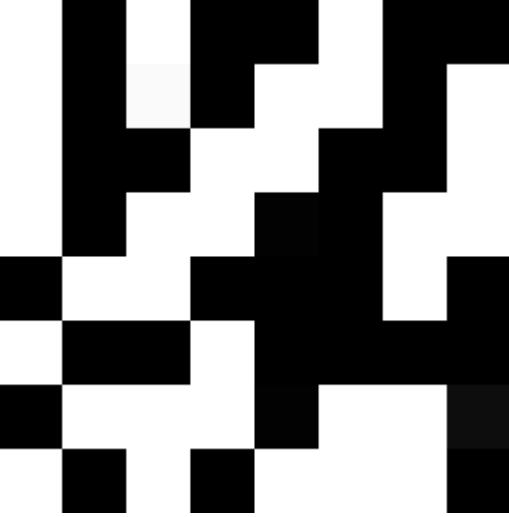
After DCT by row (normalized)



After DCT by column (normalized)



Upper left corner (highest frequency)

	<table border="1"><tr><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td></tr><tr><td>1</td><td>0</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td></tr><tr><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>0</td><td>1</td></tr><tr><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td></tr><tr><td>0</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td></tr><tr><td>1</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>0</td><td>1</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td></tr><tr><td>1</td><td>0</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td><td>0</td></tr></table>	1	0	1	0	0	1	0	0	1	0	1	0	1	1	0	1	1	0	0	1	1	0	0	1	1	0	1	1	0	0	1	1	0	1	1	0	0	0	1	0	1	0	0	1	0	0	0	0	0	1	1	1	0	1	1	0	1	0	1	0	1	1	1	0
1	0	1	0	0	1	0	0																																																										
1	0	1	0	1	1	0	1																																																										
1	0	0	1	1	0	0	1																																																										
1	0	1	1	0	0	1	1																																																										
0	1	1	0	0	0	1	0																																																										
1	0	0	1	0	0	0	0																																																										
0	1	1	1	0	1	1	0																																																										
1	0	1	0	1	1	1	0																																																										
Comparison to median	Hash 1010010010101101100110011011001101 100010100100000111011010101110																																																																

# Different File Name

## Test 1: Different File Name

Difference: different file name

Expected result: Meta-ID similar or different, Content-ID same, Data-ID same, Instance-ID same



File name: nature-PNG-1280-846-original.jpg  
 Metadata: nature-jpg-1280-846-original  
 File format: JPG  
 File size: 221 KB  
 Image size: 1280 x 846 Pixel

File name: nature-JPG-1280-846- meta-id.jpg  
 Metadata: **nature-JPG-1280-846- meta-id**  
 File format: JPG  
 File size: 221 KB  
 Image size: 1280 x 846 Pixel

Meta-ID	01001110101110101011000100111000100000111101001101101010010000101000
Content-ID	111101001101000101110001001100000100111000011111001100100110111
Data-ID	1000010000000000001010100000100001010010100011011001110000110101
Instance-ID	0001101010111000010110001100101000111110001001110011000100000110

Meta-ID	CCK4wsxLpg2MJ	CCErSovSEFGZA	88%
Content-ID	CYhx3UBzo15Le	CYhx3UBzo15Le	100%
Data-ID	CDV5N9SKqC7XN	CDV5N9SKqC7XN	100%
Instance-ID	CR5UZfzbGrnmR	CR5UZfzbGrnmR	Same

# Different File Formats

## Test 2: JPG Format vs. PNG Format

Difference: different file encoding (JPG vs. PNG)

Expected result: Meta-ID similar, Content-ID same, Data-ID different



File name: nature-PNG-1280-846-original.jpg  
 Metadata: nature-jpg-1280-846-original  
 File format: **JPG**  
 File size: 221 KB  
 Image size: 1280 x 846 Pixel

File name: nature-PNG-1280-846-original.png  
 Metadata: nature-png-1280-846-original  
 File format: **PNG**  
 File size: 1.448 KB  
 Image size: 1280 x 846 Pixel

Meta-ID	00000011100111000110000110011100011111111101110001010001101011
Content-ID	1111010011010001011100010011000000100111000011111001100100110111
Data-ID	1111100011110010101011100100011100110001110001111011100101001110
Instance-ID	010010001011110101111111011110010010001110010010000111010100101

Meta-ID	CCK4wsxLpg2MJ	CCCC2jcmEWvPK	64%
Content-ID	CYhx3UBzo15Le	CYhx3UBzo15Le	100%
Data-ID	CDV5N9SKqC7XN	CDLe7ZXaTGN1u	32%
Instance-ID	CR5UZfzbGrnmR	CRZrfwPkFvaBa	Uncorrelated

## Test 3: JPG Format vs. TIFF Format

Difference: different file encoding (JPG vs. TIFF)

Expected result: Meta-ID similar, Content-ID same, Data-ID different



File name: nature-PNG-1280-846-original.jpg  
 Metadata: nature-jpg-1280-846-original  
**File format: JPG**  
 File size: 221 KB  
 Image size: 1280 x 846 Pixel

File name: nature-TIFF-1280-846-original.png  
 Metadata: nature-TIFF-1280-846-original  
**File format: TIFF**  
 File size: 4.335 KB  
 Image size: 1280 x 846 Pixel

Meta-ID	00111011100111100000000010011100101110111100111000101110001010111
Content-ID	1111010011010001011100010011000000100111000011111001100100110111
Data-ID	111110001110000100110111001110010100011010111101100110100000111
Instance-ID	110101101101010101011111000011111100110011010111100100100000000

Meta-ID	CCK4wsxLpg2MJ	CCryAy3pj8Huc	62%
Content-ID	CYhx3UBzo15Le	CYhx3UBzo15Le	100%
Data-ID	CDV5N9SKqC7XN	CDLdmRGZDbAxq	33%
Instance-ID	CR5UZfzbGnmR	CRcw9w24r7Ct7	Uncorrelated

## Test 4: JPG Format vs. GIF Format

Difference: different file encoding (JPG vs. GIF)

Expected result: Meta-ID similar, Content-ID same, Data-ID different

	
File name: nature-PNG-1280-846-original.jpg Metadata: nature-jpg-1280-846-original <b>File format: JPG</b> File size: 221 KB Image size: 1280 x 846 Pixel	File name: nature-GIF-1280-846-original.gif Metadata: nature-GIF-1280-846-original <b>File format: GIF</b> File size: 403 KB Image size: 1280 x 846 Pixel

Meta-ID	00101011100111100000000001000110000111110011001100111010101010001101011
Content-ID	111110100110100010111000100111000000100111000011111001100100110111
Data-ID	0010011100101100111100111001111010111101010100011110000111011011
Instance-ID	0011001111101100001000011010101111100110111110110110010101001100100001

Meta-ID	CCK4wsxLpg2MJ	CC8t91o3Lg6Zv	60%
Content-ID	CYhx3UBzo15Le	CYhx3UBzo15Le	100%
Data-ID	CDV5N9SKqC7XN	CD7D43C6C4RBU	28%
Instance-ID	CR5UZfzbGrnmR	CR9gLYLAdMQti	Uncorrelated

# Different File Compression Rate

## Test 5: Compression 221 KB vs. 122 KB

Difference: different file compression rate (221 KB vs. 122 KB)

Expected result: Meta-ID similar, Content-ID same, Data-ID different



File name: nature-PNG-1280-846-original.jpg  
 Metadata: nature-JPG-1280-846-original

File format: JPG  
 File size: 221 KB  
 Image size: 1280 x 846 Pixel

File name: nature-JPG-1280-846-quality-reduction1.jpg  
 Metadata: nature-JPG-1280-846-quality-reduction1

File format: JPG  
**File size: 122 KB**  
 Image size: 1280 x 846 Pixel

Meta-ID	110001110000111010110101100111010100101101101111110101010001001010
Content-ID	1111010011010001011100010011000000100111000011111001100100110111
Data-ID	111001111101100010001000011101110010101010001001001011001101000111
Instance-ID	1001011100111111110010011001000010001001110001110011001011111110000

Meta-ID	CCK4wsxLpg2MJ	CCNt7BwGv2p6A	54%
Content-ID	CYhx3UBzo15Le	CYhx3UBzo15Le	100%
Data-ID	CDV5N9SKqC7XN	CDfn1txuBeJXg	29%
Instance-ID	CR5UZfzbGrnmR	CRmtYCtPZU5xs	Uncorrelated

## Test 6: Compression 221 KB vs. 45 KB

Difference: different file compression rate (221 KB vs. 45 KB)

Expected result: Meta-ID similar, Content-ID same, Data-ID different



File name: nature-PNG-1280-846-original.jpg  
 Metadata: nature-JPG-1280-846-original

File format: JPG  
 File size: 221 KB  
 Image size: 1280 x 846 Pixel

File name: nature-JPG-1280-846-quality-reduction2.jpg  
 Metadata: nature-JPG-1280-846-quality-reduction2  
 File format: JPG  
**File size: 45 KB**  
 Image size: 1280 x 846 Pixel

Meta-ID	11001111000001110111110101011011110101001011011011111110101010001001010
Content-ID	1111010011010001011100010011000000100111000011111001100100110111
Data-ID	10111100100111100001011110010001010100011110111000011001001001000
Instance-ID	011100000000110100101110000000101011101111001100011000000100010

Meta-ID	CCK4wsxLpg2MJ	CCbdj7yoHwLb3	56%
Content-ID	CYhx3UBzo15Le	CYhx3UBzo15Le	100%
Data-ID	CDV5N9SKqC7XN	CDKKkk3udFDbd	35%
Instance-ID	CR5UZfzbGrnmR	CRYk3TjSjEs2y	Uncorrelated

# Different Scaling

## Test 7: Scaling 1000\*661 Pixel

Difference: Scaling of JPG to 1000\*661 Pixel

Expected result: Meta-ID similar, Content-ID same, Data-ID different

	
File name: nature-PNG-1280-846-original.jpg Metadata: nature-JPG-1280-846-original File format: JPG File size: 221 KB Image size: 1280 x 846 Pixel	File name: nature-JPG-1000-661-scaling.jpg Metadata: nature-JPG-1000-661-scaling File format: JPG File size: 176 KB Image size: <b>1000 x 661 Pixel</b>

Meta-ID	010111111000111011110101011110001111010100110011111011110010110110
Content-ID	1111010011010001011100010011000000100111000011111001100100110111
Data-ID	1111000111000011001011001000001010011001010110000110111111010001
Instance-ID	110000101111110010001001100001110000010100111000111111001100001

Meta-ID	CCK4wsxLpg2MJ	CCGz382jPSmGS	47%
Content-ID	CYhx3UBzo15Le	CYhx3UBzo15Le	100%
Data-ID	CDV5N9SKqC7XN	CDhmv6Zp8sjba	35%
Instance-ID	CR5UZfzbGrnmR	CRDcgAzLJ5zxx	Uncorrelated

## Test 8: Scaling 600\*397 Pixel

Difference: Scaling of JPG to 600\*397 Pixel

Expected result: Meta-ID similar, Content-ID same, Data-ID different



File name: nature-PNG-1280-846-original.jpg  
Metadata: nature-JPG-1280-846-original  
File format: JPG  
File size: 221 KB  
Image size: 1280 x 846 Pixel

File name: nature-JPG-600-397-scaling.jpg  
Metadata: nature-JPG-600-397-scaling  
File format: JPG  
File size: 69 KB  
Image size: **600 x 397 Pixel**

Meta-ID	
Content-ID	
Data-ID	
Instance-ID	

Meta-ID	CCK4wsxLpg2MJ	CC2GEMUQcupKu -	41%
Content-ID	CYhx3UBzo15Le	CYhx3UBzo15Le -	100%
Data-ID	CDV5N9SKqC7XN	CDHbWrMKTR26i -	45%
Instance-ID	CR5UZfzbGrnmR	CRix2x5p4nuqF	Uncorrelated

# Different Croppings

## Test 9: Horizontal Cropping (Left Side)

Difference: horizontal cropping on the left side (1280 vs. 1058 px)

Expected result: Meta-ID similar, Content-ID similar or different, Data-ID different



File name: nature-PNG-1280-846-original.jpg  
 Metadata: nature-JPG-1280-846-original  
 File format: JPG  
 File size: 221 KB  
 Image size: 1280 x 846 Pixel

File name: nature-JPG-1058-846-crop-left.jpg  
 Metadata: nature-JPG-1058-846-crop-left  
 File format: JPG  
 File size: 189 KB  
 Image size: **1058** x 846 Pixel

Meta-ID	<code>1000011010011010010111011011100011000010110010011001001001001101</code>
Content-ID	<code>1111001111011100011111000011111001000010101000000110011000110001</code>
Data-ID	<code>0101001001000101000111010001110100000110111101110100000010011000</code>
Instance-ID	<code>100100010001101010010101101001111110111001110110110101100101000000</code>

Meta-ID	CCK4wsxLpg2MJ	CCVWpMK6n7VMe	45%
Content-ID	CYhx3UBzo15Le	CYhnS4yrBCCsW	42%
Data-ID	CDV5N9SKqC7XN	CDES84o7EmvXM	31%
Instance-ID	CR5UZfzbGrnmR	CRMGgy361Rn5u	Uncorrelated

## Test 10: Horizontal Cropping (Right Side)

Difference: horizontal cropping on the left side (1280 vs. 1106 px)

Expected result: Meta-ID similar, Content-ID similar or different, Data-ID different



File name: nature-PNG-1280-846-original.jpg  
 Metadata: nature-JPG-1280-846-original  
 File format: JPG  
 File size: 221 KB  
 Image size: 1280 x 846 Pixel

File name: nature-JPG-1106-846-crop-right.jpg  
 Metadata: nature-JPG-1106-846-crop-right  
 File format: JPG  
 File size: 191 KB  
 Image size: **1106** x 846 Pixel

Meta-ID	010011111000100101111001110101111110101100101111001010101101100
Content-ID	111001001101001101111001101110000001101110000011101001101100100110
Data-ID	01100100010010011111100111000001010000011010101011111110110000
Instance-ID	110101110010000111011111000111000110110110011011111110011

Meta-ID	CCK4wsxLpg2MJ	CCEiqDayzPSzb	45%
Content-ID	CYhx3UBzo15Le	CYfGu822usoYj	68%
Data-ID	CDV5N9SKqC7XN	CDTSQqRkSsA59	39%
Instance-ID	CR5UZfzbGrnmR	CRcz45B5EtQ5K	Uncorrelated

## Test 11: Vertical Cropping (Top and Bottom)

Difference: vertical cropping top and bottom (846 vs. 726 px)

Expected result: Meta-ID similar, Content-ID similar or different, Data-ID different



File name: nature-PNG-1280-846-original.jpg  
 Metadata: nature-JPG-1280-846-original

File format: JPG  
 File size: 221 KB  
 Image size: 1280 x 846 Pixel

File name: nature-JPG-1280-726-crop-top-bottom.jpg  
 Metadata: nature-JPG-1280-726-crop-top-bottom  
 File format: JPG  
 File size: 197 KB  
 Image size: 1280 x **726** Pixel

Meta-ID	11000101110010110011101101000110010111101001000100100101111
Content-ID	111101001101100101110000000111000011001101100100011101100100110110
Data-ID	0101001110111001101101110010010111111010110010101110100110100011010001
Instance-ID	1001011010011111100010000011101000001001011010011001011111000100

Meta-ID	CCK4wsxLpg2MJ	CCb44FxRQZGQN	49%
Content-ID	CYhx3UBzo15Le	CYhxA3DCmHCGR	68%
Data-ID	CDV5N9SKqC7XN	CDFCF2FS7TgtU	25%
Instance-ID	CR5UZfzbGrnmR	CRm1EKiCWfYFT	Uncorrelated

# Colour Modifications

## Test 12: Colour Modification 1

Difference: colour modification 1

Expected result: Meta-ID similar, Content-ID same or similar, Data-ID different



File name: nature-PNG-1280-846-original.jpg  
 Metadata: nature-JPG-1280-846-original

File format: JPG  
 File size: 221 KB  
 Image size: 1280 x 846 Pixel

File name: nature-JPG-1280-846-colour-modification1.jpg  
 Metadata: nature-JPG-1280-846-**colour-modification1**

File format: JPG  
 File size: 213 KB  
 Image size: 1280 x 846 Pixel

Meta-ID	<code>101011110011110011101101001110110001011110010101011010101010</code>
Content-ID	<code>1111010011010001011100010011000000100111000011111001100100110111</code>
Data-ID	<code>0011110100110100010101000000001111100010110011100110010110100010</code>
Instance-ID	<code>00100000000110111110100110010110000001010110000001101110110010</code>

Meta-ID	CCK4wsxLpg2MJ	CCWajCMfXLDys	62%
Content-ID	CYhx3UBzo15Le	CYhx3UBzo15Le	100%
Data-ID	CDV5N9SKqC7XN	CDBES3g3Qv5hH	31%
Instance-ID	CR5UZfzbGrnmR	CR6aPyYWysJ3b	Uncorrelated

## Test 13: Colour Modification 2

Difference: colour modification 2

Expected result: Meta-ID similar, Content-ID same or similar, Data-ID different



File name: nature-PNG-1280-846-original.jpg  
 Metadata: nature-JPG-1280-846-original

File format: JPG  
 File size: 221 KB  
 Image size: 1280 x 846 Pixel

File name: nature-JPG-1280-846-colour-modification2.jpg  
 Metadata: nature-JPG-1280-846-**colour-modification2**  
 File format: JPG  
 File size: 220 KB  
 Image size: 1280 x 846 Pixel

Meta-ID	1010111010011110101111010001110110000011110011111001010101101010
Content-ID	1111011011010001110000010011000011100111111001000101100100100110
Data-ID	100100010101111110001100001101101100101011101000001001011000111
Instance-ID	0001001110010100010110011111010100100010110110100100101010101100

Meta-ID	CCK4wsxLpg2MJ	CCWZ4bsnaouRH	58%
Content-ID	CYhx3UBzo15Le	CYLTHW4TGFjt5	60%
Data-ID	CDV5N9SKqC7XN	CDMYY3YWz6SWE	27%
Instance-ID	CR5UZfzbGrnmR	CR4GwqBi7mZyd	Uncorrelated

## Test 14: Colour Modification 3

Difference: colour modification 3

Expected result: Meta-ID similar, Content-ID same or similar, Data-ID different



File name: nature-PNG-1280-846-original.jpg  
 Metadata: nature-JPG-1280-846-original

File format: JPG  
 File size: 221 KB  
 Image size: 1280 x 846 Pixel

File name: nature-JPG-1280-846-colour-modification3.jpg  
 Metadata: nature-JPG-1280-846-**colour-modification3**  
 File format: JPG  
 File size: 91 KB  
 Image size: 1280 x 846 Pixel

Meta-ID	101011110011111001110110100111011000000111100111100001010101001010
Content-ID	111100001111000001111000000000111100001111100011110011001111001100110011
Data-ID	111100000010001000001110010101011101010100010100111001110011101101
Instance-ID	11110001000111110000110011110001100011001101000110010001110101010101

Meta-ID	CCK4wsxLpg2MJ	CCWY6Ct7GKXF	56%
Content-ID	CYhx3UBzo15Le	CYhtvuuWzEZKE	68%
Data-ID	CDV5N9SKqC7XN	CDhrbTpAzHX18	41%
Instance-ID	CR5UZfzbGrnmR	CRhiBALBatbzk	Uncorrelated

## Test 15: Colour Modification 4 (black & white)

Difference: colour Modification 4 (black & white)

Expected result: Meta-ID similar, Content-ID same or similar, Data-ID different



File name: nature-PNG-1280-846-original.jpg  
 Metadata: nature-JPG-1280-846-original

File format: JPG  
 File size: 221 KB  
 Image size: 1280 x 846 Pixel

File name: nature-JPG-1280-846-black-white.jpg  
 Metadata: nature-JPG-1280-846-**black-white**  
 File format: JPG  
 File size: 201 KB  
 Image size: 1280 x 846 Pixel

Meta-ID	10001111110110100011010010111100111000111111111110101010101101000
Content-ID	1111010011010001011100010011000000100111000011111001100100110111
Data-ID	11000000111100110101011011110000010100100000011101010111001001
Instance-ID	0110111100110010000010110010010100000110010010100001111011010001

Meta-ID	CCK4wsxLpg2MJ	CCM4KoCAkcUzB	62%
Content-ID	CYhx3UBzo15Le	CYhx3UBzo15Le	100%
Data-ID	CDV5N9SKqC7XN	CDDGqzjVCj36G	27%
Instance-ID	CR5UZfzbGrnmR	CRYbjf8C4uG8G	Uncorrelated

# Image Manipulation

## Test 16: Flock of Birds Removed

Difference: image manipulation, the flock of birds has been removed

Expected result: Meta-ID similar, Content-ID same or similar, Data-ID different



File name: nature-PNG-1280-846-original.jpg  
 Metadata: nature-JPG-1280-846-original  
 File format: JPG  
 File size: 221 KB  
 Image size: 1280 x 846 Pixel

File name: nature-JPG-1280-846-nobirds.jpg  
 Metadata: nature-JPG-1280-846-**nobirds**  
 File format: JPG  
 File size: 150 KB  
 Image size: 1280 x 846 Pixel

Meta-ID	011011111010110011111000101110011110101011010111110101010001101010
Content-ID	1111010011010001011100010011000000100111000011111001100100110111
Data-ID	10010010101110010110110110111011000111000010010101010001000111010
Instance-ID	011001010101011100010110000100001011011101100011011110001101110001101

Meta-ID	CCK4wsxLpg2MJ	CCYgVq465DG4C	58%
Content-ID	CYhx3UBzo15Le	CYhx3UBzo15Le	100%
Data-ID	CDV5N9SKqC7XN	CDMKrYqib1wNR	28%
Instance-ID	CR5UZfzbGrnmR	CRTx98fu6zpCz	Uncorrelated

## Test 17: Flock of Birds and Tree Removed

Difference: image manipulation, the flock of birds has been removed, one tree modified

Expected result: Meta-ID similar, Content-ID similar, Data-ID different



File name: nature-PNG-1280-846-original.jpg  
 Metadata: nature-JPG-1280-846-original

File format: JPG  
 File size: 221 KB  
 Image size: 1280 x 846 Pixel

File name: nature-JPG-1280-846-nobirds-notree.jpg  
 Metadata: nature-JPG-1280-846-nobirds-notree  
 File format: JPG  
 File size: 954 KB  
 Image size: 1280 x 846 Pixel

Meta-ID	0010111100010001011100010011001111000101101011111010111011010100
Content-ID	1111011011011001001011110000010001000100111000111001100100100100111
Data-ID	110110010101101001000110001000000110000011001111010111101111110111
Instance-ID	11001101100111101100101111000011011111000111000111110001010011101

Meta-ID	CCK4wsxLpg2MJ	CC8x97wCgMM2X	60%
Content-ID	CYhx3UBzo15Le	CYLTKHXrsxJ62	68%
Data-ID	CDV5N9SKqC7XN	CDdAbVBuuiQD1	38%
Instance-ID	CR5UZfzbGrnmR	CRbVSQFr8WJut	Uncorrelated

## Test 18: Water vs. Grass

Difference: image manipulation, water vs. grass

Expected result: Meta-ID similar, Content-ID similar, Data-ID different



File name: nature-PNG-1280-846-original.jpg  
 Metadata: nature-JPG-1280-846-original

File format: JPG  
 File size: 221 KB  
 Image size: 1280 x 846 Pixel

File name: nature-JPG-1280-846-trees-on-grass.jpg  
 Metadata: nature-JPG-1280-846-trees-on-grass  
 File format: JPG  
 File size: 265 KB  
 Image size: 1280 x 846 Pixel

Meta-ID	00101110000010101011010110011001110000111101001111010000000101010
Content-ID	1101010011010001000110010010011000100110110010011111001100101110011
Data-ID	1010000111011110101000001000010111110000001010010110010001101111
Instance-ID	01101100110100101101100111011110110000100100011011100101100100

Meta-ID	CCK4wsxLpg2MJ	CC8hfUsJ6cJdo	56%
Content-ID	CYhx3UBzo15Le	CYcbbBuVquKUr	68%
Data-ID	CDV5N9SKqC7XN	CDU5icqHNxabK	35%
Instance-ID	CR5UZfzbGrnmR	CRY1j7m6qcUJb	Uncorrelated

## Test 19: One Pixel Change JPG

Difference: image manipulation, one pixel change in a JPG file

Expected result: Meta-ID similar, Content-ID same, Data-ID similar or different

	
File name: nature-PNG-1280-846-original.jpg Metadata: nature-JPG-1280-846-original File format: JPG File size: 221 KB Image size: 1280 x 846 Pixel	File name: nature-JPG-1280-846-one-pix.jpg Metadata: nature-JPG-1280-846-one-pix File format: JPG File size: 221 KB Image size: 1280 x 846 Pixel

Meta-ID	101011111001101011110100101101110101100010110001111001010001101010
Content-ID	1111010011010001011100010011000000100111000011111001100100110111
Data-ID	1100010111100110010001101000100001001011101000011011000111001001
Instance-ID	1011010000011010000011110010100101110001000110100100110100010001

Meta-ID	CCK4wsxLpg2MJ	CCWabttwqAFyY	73%
Content-ID	CYhx3UBzo15Le	CYhx3UBzo15Le	100%
Data-ID	CDV5N9SKqC7XN	CDN6seVih2GKK	39%
Instance-ID	CR5UZfzbGrnmR	CRX8ZAKReu89a	Uncorrelated

## Test 20: One Pixel Change GIF

Difference: image manipulation, one pixel change in a GIF file

Expected result: Meta-ID similar, Content-ID same, Data-ID similar or different

	
File name: nature-GIF-1280-846-original.gif Metadata: nature-GIF-1280-846-original File format: GIF File size: 403 KB Image size: 1280 x 846 Pixel	File name: nature-GIF-1280-846-one-pix.gif Metadata: nature-GIF-1280-846-one-pix.gif File format: GIF File size: 403 KB Image size: 1280 x 846 Pixel

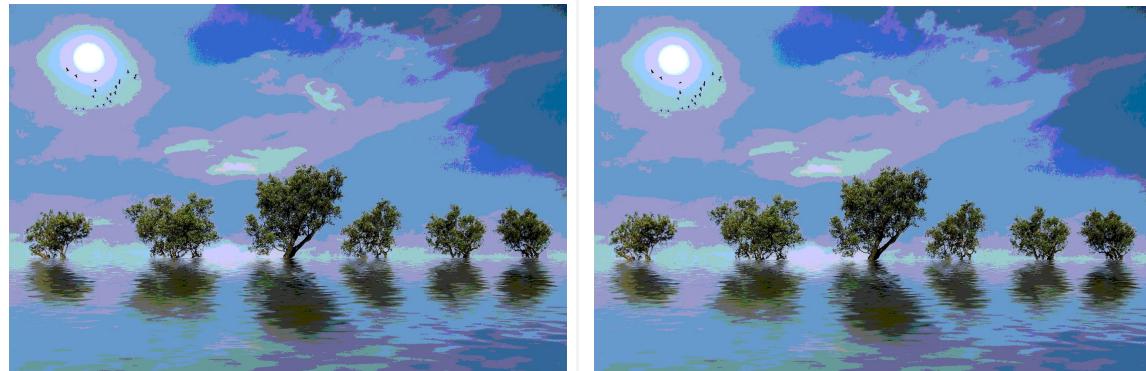
Meta-ID	0010101110001010001011001010110100110010011001100010001010001111110
Content-ID	11110100110100010111000100110000001001110000111110011001001101111
Data-ID	001001110011110011100111001011011111000101000101001011101101111
Instance-ID	0001000110110101111000101001100001111001111100000111111000010000100

Meta-ID	CC8t91o3Lg6Zv	CC8TVdrBvtXAf	64%
Content-ID	CYhx3UBzo15Le	CYhx3UBzo15Le	100%
Data-ID	CD7D43C6C4RBU	CD7Df4L2r13fk	80%
Instance-ID	CR9gLYLAdMQti	CR4FvhHjeDSYR	Uncorrelated

## Test 21: One Pixel Change GIF

Difference: image manipulation, one pixel change in a GIF file (saved with web-save color palette)

Expected result: Meta-ID similar, Content-ID same, Data-ID similar or different



File name: nature-GIF-1280-846-fxc.gif

Metadata: nature-GIF-1280-846-fxc

File format: GIF

File size: 127 KB

Image size: 1280 x 846 Pixel

File name: nature-GIF-1280-846-fxc-one-pix.gif

Metadata: nature-GIF-1280-846-fxc-one-pix

File format: GIF

File size: 127 KB

Image size: 1280 x 846 Pixel

Meta-ID 0110101111101000001011001010101101010001101110

Content-ID 111101101101000101110001001010000100111000011110100100100100110111

Data-ID 1100110010110110010101010011010101011001010100110100000000011

Instance-ID 0100111010011010101000101011001100111011010111011010010011110000

Meta-ID	CCY2cxaWr2UVF	CCY3RTZfXuWoo	75%
Content-ID	CYLTmp1L5QD2a	CYLTmp1L5QD2a	100%
Data-ID	CDbExyhmryY84	CdbExyhmryY86	97%
Instance-ID	CRGtvuxoXvv6M	CRE9DA79jvKXT	Uncorrelated

## Test 22: One Pixel Change BMP

Difference: image manipulation, one pixel change in a BMP file

Expected result: Meta-ID similar, Content-ID same, Data-ID similar or different



File name: nature-BMP-1280-846-original.bmp  
 Metadata: nature-BMP-1280-846-original  
 File format: BMP  
 File size: 3.248 KB  
 Image size: 1280 x 846 Pixel

File name: nature-BMP-1280-846-original.bmp  
 Metadata: nature-BMP-1280-846-original  
 File format: BMP  
 File size: 3.248 KB  
 Image size: 1280 x 846 Pixel

Meta-ID	<code>001001010000100001111001010110100101010110001100000010101101000</code>
Content-ID	<code>1111010011010001011100010011000000100111000011111001100100110111</code>
Data-ID	<code>00010001111101010111101110000101100001001001101111001100110000011</code>
Instance-ID	<code>101110101110101100110111011011011010110001110101111111000000110</code>

Meta-ID	CCCB2ZwZ27cPe - CYhx3UBzo15Le - CD4CZzUwKoeYU - CRbozQjJB9Hmr	CC71GKmSpRvsh CYhx3UBzo15Le CD4CZzPrdstc2 CRKGA77NGSPcy	71% 100% 97% Uncorrelated
---------	--	--	------------------------------------

## Test 23: Visible Watermark in JPG

Difference: image manipulation, watermark in a JPG file

Expected result: Meta-ID similar, Content-ID same or similar, Data-ID different



File name: nature-PNG-1280-846-original.jpg  
 Metadata: nature-JPG-1280-846-original

File format: JPG  
 File size: 221 KB  
 Image size: 1280 x 846 Pixel

File name: nature-JPG-1280-846-watermark.jpg  
 Metadata: nature-JPG-1280-846-watermark  
 File format: JPG  
 File size: 268 KB  
 Image size: 1280 x 846 Pixel

Meta-ID	
Content-ID	
Data-ID	
Instance-ID	

Meta-ID	CCK4wsxLpg2MJ	CCbjBH9QxuV8Z -	86%
Content-ID	CYhx3UBzo15Le	CYhx3UB6mwmNm -	94%
Data-ID	CDV5N9SKqC7XN	CD3u6ZMScDzq9 -	33%
Instance-ID	CR5UZfzbGrnmR	CRLGwias6xegj	Uncorrelated

# Reproduction

## Test 24: Photo of the image

Difference: photo taken from the image on the desktop computer with a smartphone

Expected result: Meta-ID similar, Content-ID same or similar, Data-ID different

	
File name: nature-JPG-1280-846-original.jpg Metadata: nature-JPG-1280-846-original File format: JPG File size: 221 KB Image size: 1280 x 846 Pixel	File name: nature-JPG-3594-2603-photo.jpg Metadata: nature-JPG-3594-2603- <b>photo</b> File format: JPG File size: 1.528 KB Image size: 3594 x 2603 Pixel

Meta-ID	0010101110111100110100001001100011100111011111111011000000000010
Content-ID	11010100110110010011110010010010111000100111000011110001100100110110
Data-ID	01110110111101100000000001010011011010111000010010001000101110011
Instance-ID	100000011100100000100101000001101001010110000000011010010101110001000

Meta-ID	CCK4wsxLpg2MJ	CC8YtQp76dzK9	58%
Content-ID	CYhx3UBzo15Le	CYcbu3hPccNny	73%
Data-ID	CDV5N9SKqC7XN	CDiu55BBVUaEQ	35%
Instance-ID	CR5UZfzbGrnmR	CRVC21vBzbchS	Uncorrelated

# Results

## Meta-ID

For all tests the file name was used as an input for the Meta-ID. Due to the input being relatively short, changes of only a few characters result in a relatively significant change of the Meta-ID.

## Content-ID-Image

**The Content-ID shows a strong resilience to changes of format, compression, scaling, color modifications or small image manipulation!**

The cropping of an image can have a significant influence on the calculation of the hash and the degree of similarity between two images.

One unexpected result: [Horizontal cropping \(left side\)](#) results in no significant similarity while [Horizontal cropping \(right side\)](#) does. The difference in similarity when comparing left-side vs right-side cropping shows that cropping operations result in variable amounts of changes in the Content-ID-Image depending on the structure of the image.

## Data-ID

The Data-ID shows that even minor changes of compression, scaling, color modifications or small image manipulation results in a loss of any significant data similarity.

But as the tests 19 to 22 show: comparing two images with a change of only 1 single pixel will result in a different similarity of the Data-ID which is depending on the encoding per image file format.

## Instance-ID

As expected, the Instance-ID, which is built from the raw data of the media object using a SHA256-based hash-tree, is entirely uncorrelated in all tests.

Please feel free to send us your feedback, findings and other comments.

# ISCC Batch Result

CCaXpicfe6UJB	CYLTmp1L5QD2a	CDbExyhmryY86	CRE9DA79jvKXT	4e9aa2b33b5da4f06 e6b76548ff0719ec30 90242bbe1f5b8b427 90858b2d7b8c	nature-GIF-1280-846 -fxc-one-pix.gif
CC7na3ep5Vv4p	CYLTkHXrsxJ62	CDdAbVBuuiQD1	CRbVSQFr8WJut	cd9ecbc37c71f29d88 64e7d15337e93b644 b4bcf503812b7e198 bc0e7619f38c	nature-JPG-1280-846-nobirds-notree.jpg
CCErSovSEFGZA	CYhx3UBzo15Le	CDV5N9SKqC7XN	CR5UZfzbGrnmR	1ab858ca3e2731062 f8936a948e31532b0 b8abc2f5389122e04 d717416924f6b	nature-JPG-1280-846-meta-id.jpg
CCDShqEPAkknk2	CYhx3UBzo15Le	CDKKkk3udFDbd	CRYk3TjSjEs2y	700d2c0adef318229 b1661d970404ca05f b6346e55d01f611a1 599bfa71e6c54	nature-JPG-1280-846-quality-reduction2.jpg
CCv1R8yoZfJmz	CYLTmp1L5QD2a	CDbExyhmryY84	CRGtvuxoXvv6M	5b795775a5ad9a6e3 7b5e10c37d46178f3 b98990c185bb5fb9e d32265ffd055f	nature-GIF-1280-846-fxc.gif
CCV4pFBECsRrt	CYhx3UBzo15Le	CDfn1txuBeJXg	CRmtY CtPZU5xs	973fe4c84e3997f07a ed68d87f233faf12e1 52dcf6a702398ca1fb a48de063fb	nature-JPG-1280-846-quality-reduction1.jpg
CCmo1PLFQpfjsj	CYhx3UBzo15Le	CD7D43C6C4RBU	CR9gLYLAdMQti	33ec2357cdf65321e2 6623fe23d7616fe0d9 4d716794d8e0f14fb2 c90fff7705	nature-GIF-1280-846-original.gif
CCK4wsxLpg2MJ	CYhx3UBzo15Le	CDV5N9SKqC7XN	CR5UZfzbGrnmR	1ab858ca3e2731062 f8936a948e31532b0 b8abc2f5389122e04 d717416924f6b	nature-JPG-1280-846-original.jpg
CCmJ4EcgADoRx	CYhx3UBzo15Le	CD4CZzUwKoeYU	CRbozQjJB9Hmr	d01e08e4c84cbe1b8 b1bc2817b71567af7 5835212ed696d2043 c4868e23b4b00	nature-BMP-1280-846-original.bmp
CCasZvx1tVeT9	CYhx3UBzo15Le	CD7Df4L2r13fk	CR4FvhHjeDSYR	136bca61e7c1f0842b 800962994c8ebe451 b5ec9a69a773edd1a 754385b16458	nature-GIF-1280-846-one-pix.gif
CCTkiKDqh8ke	CYhnS4yrBCCsW	CDES84o7EmvXM	CRMGgy361Rn5u	911a95a7f73b5940b c3c8cead336f311667 e67db9b96cc78771a e6c084e31e4c	nature-JPG-1058-846-crop-left.jpg
CC7x5x7aFsdj4	CYhx3UB6mwmNm	CD3u6ZMScDzq9	CRLGwias6xegj	f6c407927715b31c03 0813e085ccb86f2b9f 2d780b348afdf0629 27699978a24	nature-JPG-1280-846-watermark.jpg
CCvAUYSCh4tbr	CYhx3UBzo15Le	CDBES3g3Qv5hH	CR6aPyYWysJ3b	201bf4cb02b037b2fa 49426d28fad690ea9 c7e41ab017bc25127f	nature-JPG-1280-846-colour-modification1.jpg

				440ea0a7b4c	
CCvAUUiCHVdETQ	CYLTHW4TGFjt5	CDMYY3YWz6SWE	CR4GwqBi7mZyd	139459f522da4aac2 ee830def179dcf4c4b e727ada1955b6c6c1 0803c772fa40	nature-JPG-1280-84 6-colour-modification 2.jpg
CCYCwRcv9oF9Y	CYhxA3DCmHCGR	CDFCF2FS7TgtU	CRm1EKiCWfYFT	969f883a096997c4a 0265d811d0c83d2af 20f6b8fb2e9296174 26bbbebad5f8	nature-JPG-1280-72 6-crop-top-bottom.jpg
CCv24yyRj7iYm	CYhtvuuWzEZkE	CDhrbTpAzHX18	CRhiBALBatbzk	f11f0cf8cd191d5591c 441ff46d0946b01b0a eb122ea20b2171c88 b6d63e75d6	nature-JPG-1280-84 6-colour-modification 3.jpg
CCPgFncCMY5aa	CTrQNaP2oievx	CDBQJXgzSmin3	CR4HrfLfgp6c4	14a21142b9f8f26d7b bd7f1ea2d7814de36 8d40acc3521df1bb33 1417e081cb5	README.md
CC7yAmWiEwmgp	CYhx3UBzo15Le	CDMKrYqib1wNR	CRTx98fu6zpCz	655771610b76378d8 9a99d153434705218 20546074672ca6ce4 5431c10f2e0e6	nature-JPG-1280-84 6-nobirds.jpg
CC2dX5AJpcscp	CYhx3UBzo15Le	CDN6seVih2GKK	CRX8ZAKReu89a	b41a0f29711a4d117 0d8019bd5c5703ddc 97cebf2045aaed913 b05aaafacfd83	nature-JPG-1280-84 6-one-pix.jpg
CC7hCmKhn7xGL	CYfGu822usoYj	CDTSQqRkSsA59	CRcz45B5EtQ5K	d721df8e36cd97f348 d52fc076df68e690c2 71ecfd70eef73e459e 98abace0ab	nature-JPG-1106-84 6-crop-right.jpg
CCBLGWs3B3bDQ	CYcbbBuVquKUr	CDU5icqHNxabK	CRY1j7m6qcUJb	6cd2dcf091b9646b 1bbdcf8c13e74fc534 a08e1aa3025331e91 327b344ce37	nature-JPG-1280-84 6-trees-on-grass.jpg
CCKEe9NyMvhKt	CYhx3UBzo15Le	CDLe7ZXaTGN1u	CRZrfwPkFvaBa	48bd7fdfe48c90ea52c ad8622161c1fe5c351 152c82db6d9a7cf4c3 1fdf28973a	nature-PNG-1280-84 6-original.png
CCtAf5eKBJSmy	CYcbu3hPccNny	CDiu55BBVUaEQ	CRVC21vBzbchS	83881286b00d2b884 dc81e7627b9cced67 549d83b4b529a0659 a866a429f71a5	nature-JPG-3594-26 03-photo.jpg
CC99JMXj3ebgg	CYhx3UBzo15Le	CDtVCLvNUM2HJ	CRtMU2M9YF91y	6828860a30b481a6e 603100acf1ecfb9bfd 3d91e90e9eafc1782 e10e281b0c4	nature-PNG-1280-84 6-nobirds.png
CCcFA9qTN6f12	CYhx3UBzo15Le	CD4CZZPrdstc2	CRKGAT7NGSPcy	baeb376dac75ff06dff e6c6cef159447e66a7 519605606f9937aab a37d4ddb1	nature-BMP-1280-84 6-one-pix.bmp
CC5xi54G8d2Cm	CYhx3UBzo15Le	CDDGqzjVCj36G	CRYbjf8C4uG8G	6f320b25064a1ed1f2 8b9ebae8221fd118c bd5a1d175882a1b3b 01641f6ece9a	nature-JPG-1280-84 6-black-white.jpg