



# About Photo Metadata

Considerations and advice shared by  
**Michael Steidl**  
Lead of IPTC Photo Metadata Working Group  
at the IPTC Photo Metadata Conference  
13 October 2020



# Who is the PMD Working Group

A group of photo metadata people from an IPTC member (in random order)



David Riecks, PLUS



Karl Csoknyay, Keystone-SDA



Brendan Quinn, IPTC MD



Michael Steidl (Lead)



Annette Feldman, AP



Linda Burman



Jeff Sedlik, PLUS



Alfonso Gutierrez, CEPIC



Margaret Warren



Scott Eells, Bloomberg



Alison Sullivan



## Let's talk about ...

### Photo Metadata (PMD)

MUST  
should  
can  
...  
may  
does not

} tell the truth

3



## Clarification

Photo Metadata telling the truth

=

True and correct facts  
about the image

(... and we don't talk about the truth of the pixels)

4



## The Metadata Sea and its Photo Metadata Bay



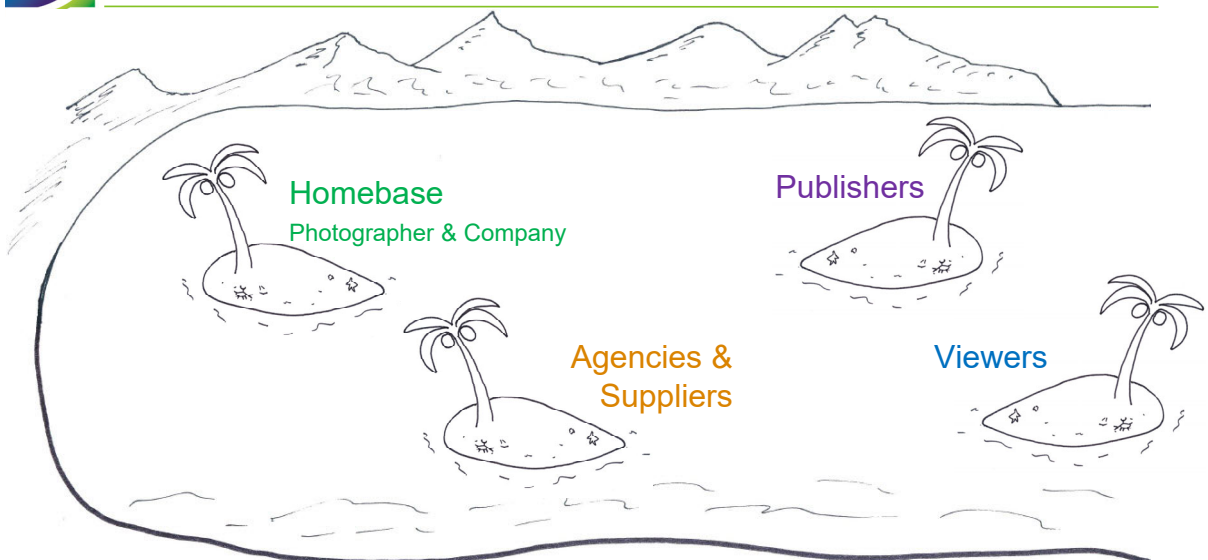
© 2020 IPTC (www.iptc.org) All rights reserved

5

5



## The Islands in the Photo Metadata Bay



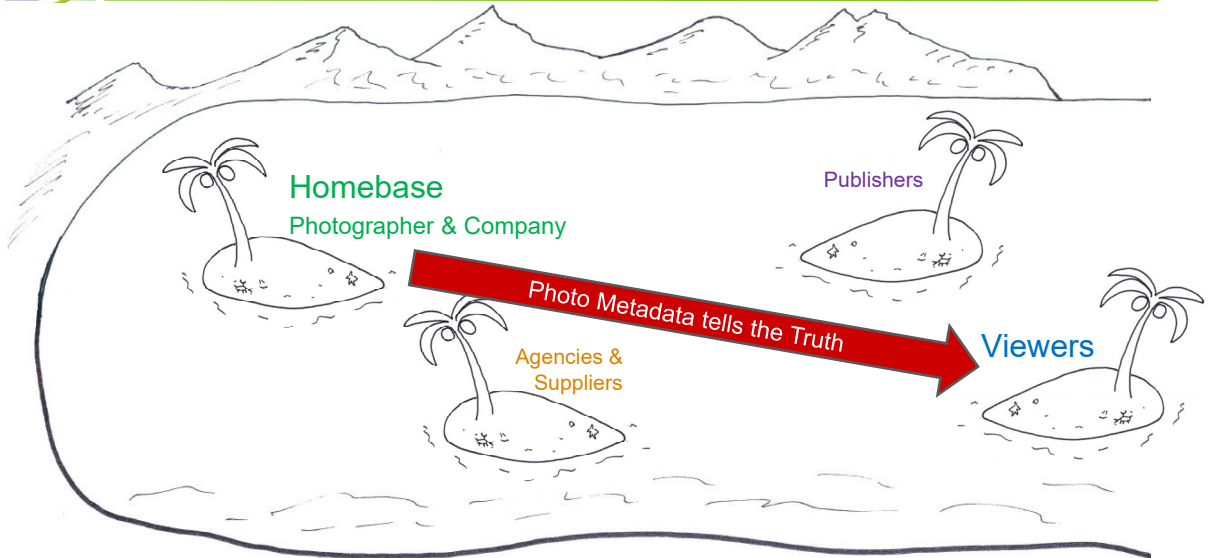
© 2020 IPTC (www.iptc.org) All rights reserved

6

6



## The Goal of Photo Metadata: Communication



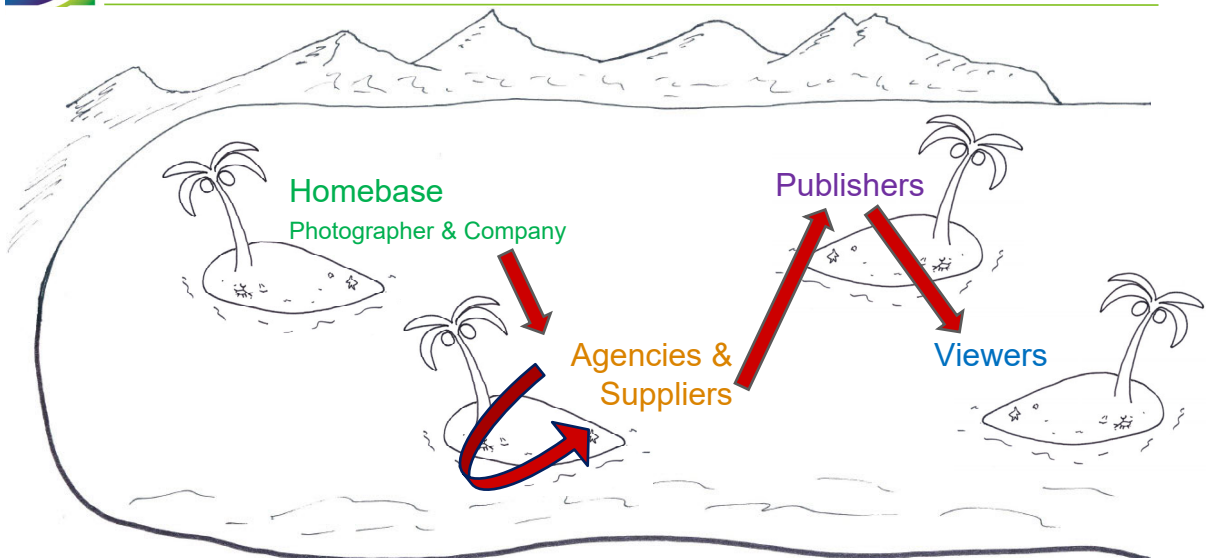
© 2020 IPTC (www.iptc.org) All rights reserved

7

7



## PMD communication reality check

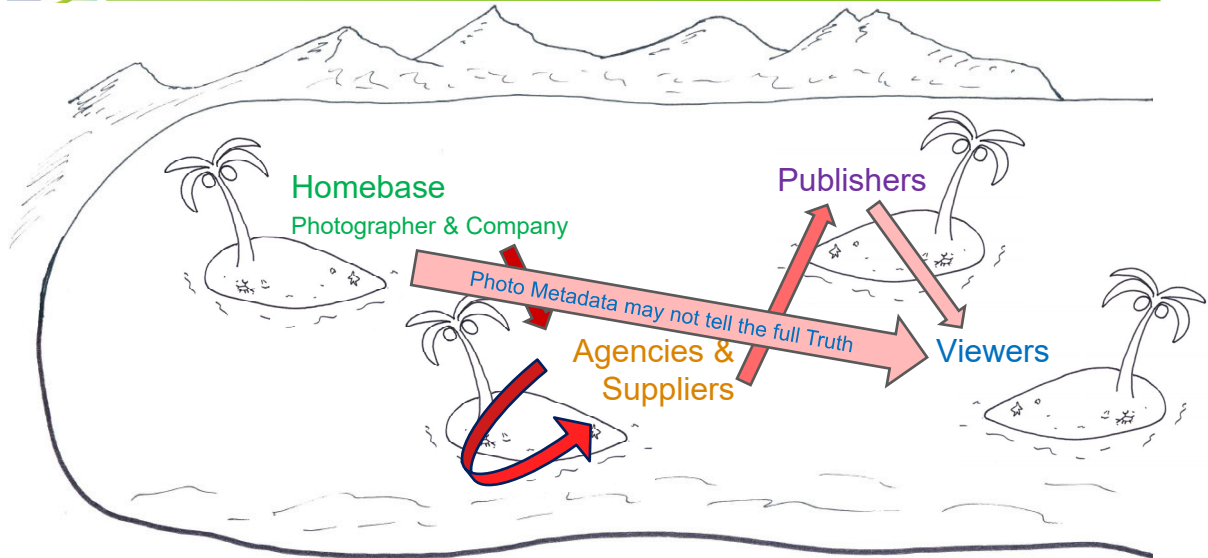


© 2020 IPTC (www.iptc.org) All rights reserved

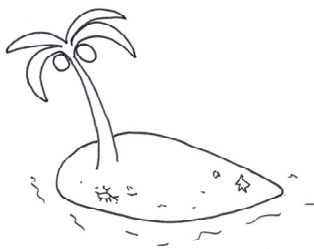
8

8

## PMD's fate in the communication



## Homebase island



= the creator(s) of the photo, the photographer and her/his company

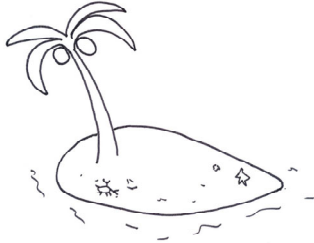
Metadata is based on knowledge very close to the pictured event or object, it is checked and written down.

### Other involved parties:

- Camera makers
- Makers of imaging software supporting photo metadata standards (e.g. Lightroom, Photo Mechanic, Fotostation, Photoshop ...)
- Makers of Digital Asset Management software supporting photo metadata standards
- Providers of Artificial Intelligence/Machine Learning systems supporting images



## Agencies/Suppliers island



= companies collecting, aggregating, supplying and licensing photos

Metadata may be edited: let's assume corrected or metadata required for attracting customers (e.g. keywords) added.

**Other involved parties:**

- Makers of imaging software supporting photo metadata standards
- Makers of DAM software supporting photo metadata standards
- AI/ML system providers supporting images



## Publishers island



= organisation posting licensed photo(s) on the web (or in print)

Metadata is used, replaced or removed as the organisation is interested in.

**Other involved parties:**

- Makers of editorial systems for journalists
- Web Content Management system makers
- Content Delivery Networks reformatting photos
- Search engines crawling the website and showing found images in search results ... with metadata



## Viewers island



= persons viewing a photo and reading metadata information (if readable)

Metadata must be displayed to them – else they can't get the metadata information.

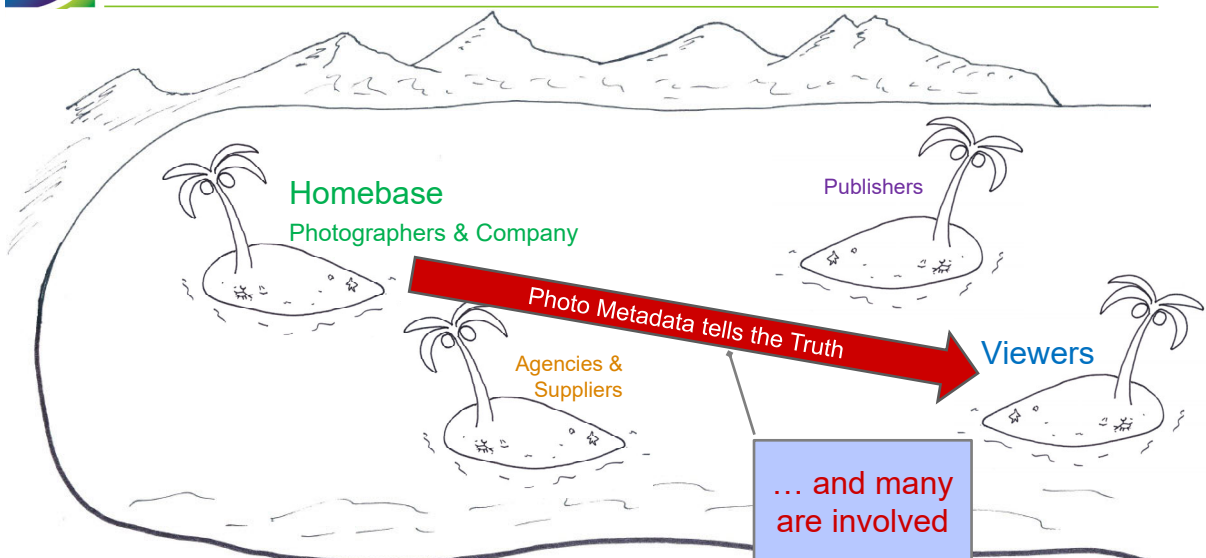
### Other involved parties:

- Makers of software showing the photo:  
Image viewers, web browsers, PDF viewers, file metadata viewers
- Search engines showing photos as results – let's hope - with metadata

13



## The Goal of Photo Metadata: Communication



14



## How to communicate photo metadata well

- This communication needs a **language** – beyond single words
- A language needs a grammar, rules for building sentences making a statement:

Most languages of Europe and the Americas use this basic grammar ...

Subject	Predicate	Object
“Frank”	“takes”	“a photo”
“Francois”	“prend”	“une photo”
“Franz”	“schießt”	“ein Foto”
“Francesco”	“toma”	“una foto”
“Franco”	“scatta”	“una foto”
“Frank”	“robi”	“zdjęcie”

© 2020 IPTC (www.iptc.org) All rights reserved

15

15



## How to communicate photo metadata well

Many standards for photo metadata follow this basic grammar and build statements this way:

- The **Subject** is “**this image**”
- The **Predicate** is a field/property/tag **defined by a (photo) metadata standard**  
Such a predicate defines what this statement tells about the photo!
- The **Object** is the value of a field **set by a metadata editor**

Example: Subject=“**This image**”, Predicate=“**has Creator**”, Object=“**Carola Ranotti**”

The field label ↑

The value of the field ↑



Photo metadata fields tell simple sentences, make statements for the human viewer.

© 2020 IPTC (www.iptc.org) All rights reserved

16

16





## Dear Photo Metadata Standard User ...

... take care to use a metadata field as defined by its standard

This supports

For IPTC Photo Metadata this helps:

- [The User Guide](#), grouping fields by these categories
  - [General image content](#)
  - [Person\(s\) shown in the image](#)
  - [Locations relevant to the image](#)
  - [Other things shown in the image, including artwork](#)
  - [How to assert rights and license usage](#)
  - [Administration and Commissioning details](#)
- ... and giving hints for their use.

The [Field Reference Table](#) lists all IPTC fields with their field labels for easy reference.

- [The Standard document](#) with a definition of each field (= property)

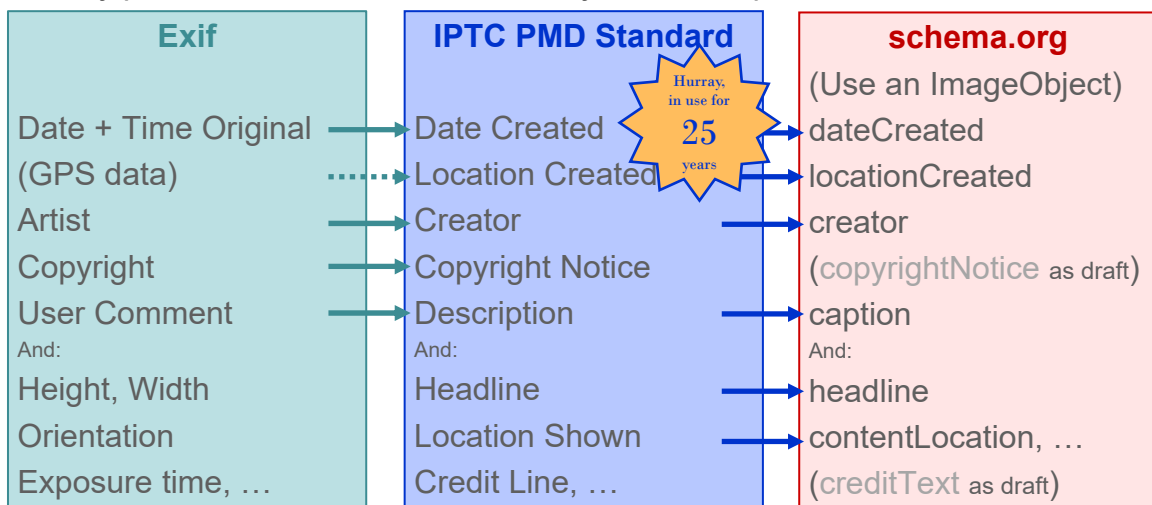
E.g.:

Name	Creator
Definition	Contains the name of the photographer, but in cases where the photographer should not be identified the name of a company or organisation may be appropriate.



## Photo Metadata Standards out there

... they provide a well defined vocabulary of fields = predicates of communication



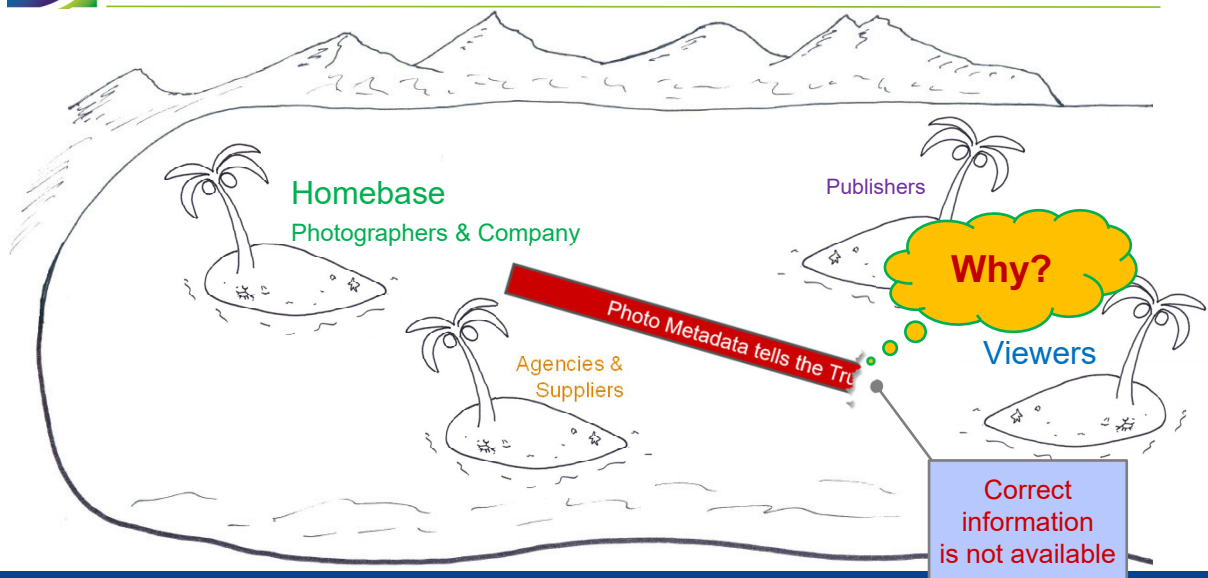


## Use of standards in a photo workflow

	Metadata presets	Taking the photo	Edit metadata /1	Edit metadata /2...	Photo in web page
	Metadata default values are set in the camera by its user. Preset of metadata values in a file on the memory card.	Metadata of the camera and of a preset on the memory card is embedded into the image file.	Exif Tag values are copied to IPTC fields. Metadata fields are edited, IPTC fields and some Exif Tags are updated & embedded.	Metadata fields are edited, IPTC fields and some Exif Tags are updated synchronously and embedded.	Exif and IPTC metadata should stay embedded, its values can be copied to schema.org metadata properties.
Exif	Exif Tags can be set	Exif data is generated and embedded	Exif data is copied to IPTC; may be updated	Exif data is updated (in sync with IPTC)	Exif data should not be removed
IPTC	IPTC fields may be set	IPTC presets can be copied and embedded	IPTC fields may get copy of Exif values; is updated	IPTC data is updated (in sync with Exif)	IPTC data should not be removed
schema.org					If schema.org is used: IPTC and Exif values should be copied to Image Object properties



## The threat: broken communication





## Why? → Check is the input correct?

- ✓ Photographer when taking the photo: is **date and time of the camera correct**, **can GPS data be received** properly?
- ✓ Photographer as first PMD editor: are all known **data written down correctly**?
- ✓ Photo metadata editor on a desk: **are all values written by the photographer correct** including spelling? (4-eyes principle)  
If a geo-location service is used with the GPS data:  
do the generated location names match information of the photographer?
- ✓ **Keyword/term added by Artificial Intelligence (AI) or Machine Learning (ML): does the provided term match** the content closely? (The “score” helps.)



## Why? → Check are the values kept alive?

Keeping alive = removed only with good reasons and not removed inadvertently

- ✓ **Check what metadata fields and values are embedded** into an image file
  - ⇒ Imaging software with photo metadata support can be used – but some pieces of software don't **support all IPTC fields**.
  - ⇒ IPTC suggests to use its **Get IPTC Photo Metadata**: <https://getpmd.iptc.org/>
- ✓ Check if **metadata fields displayed and used by Google Images are embedded**
- ✓ **Check if embedded metadata of a photo posted on the web exists** as expected:
  - Check your own website, in your web shop
  - Check the website of your customers
  - (See what the [IPTC Photo Metadata Explorer](#) found on widely used news sites)

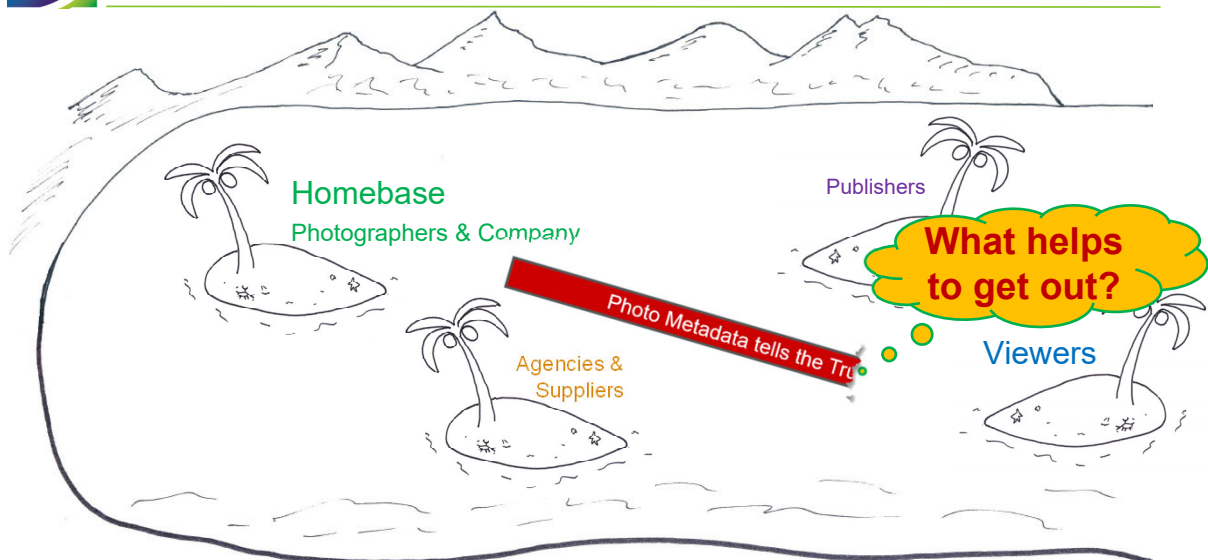


## Why? → Check are the values kept alive? ... continued

- ✓ Check how open systems you are using for displaying photos treat photo metadata:  
IPTC has checked the fate of embedded metadata on 500px, Behance, EyeEm, Flickr, fotocommunity.de, Google Photos, Instagram and more in 2019, and [published the results and how the sites were tested](#).
- ✓ Check your license contracts: do they tell customers to keep metadata alive? (... at least fields relevant for rights ....)



## The threat: broken communication





## Helpers for getting out (1)

- The Photo Metadata Bay needs a guide for synchronizing metadata values across standards; also called mapping of fields.

Currently IPTC discusses this goal with CIPA, the authority behind Exif, and Schema.org

Preview snippet →

Property Name	Definition	Exif Tags (as of: 2020.08.14) Shown as: Tag Name / Field Name (Tag ID Dec, Hex) Ref. DC-008-2019-E=Exif.pdf (v 2.32)	Schema.org (as of: 2020-09-30 IPTC PMD WG) Use Image Object <a href="https://schema.org/ImageObject">https://schema.org/ImageObject</a> and its properties listed below XXX = no mapping
Creator	Contains the name of the photographer, but in cases where the photographer should not be identified the name of a company or organisation may be appropriate.	Person who created the image / Artist (315, 0x013B)	<a href="https://schema.org/creator">https://schema.org/creator</a> (see <a href="https://schema.org/author">https://schema.org/author</a> too if required by HTML5 purposes)
Creator's Contact Info	The creator's contact information provides all necessary information to get in contact with the creator of this image and comprises a set of sub-properties for proper addressing.		<a href="https://schema.org/creator-&gt;">https://schema.org/creator-&gt;</a> (Person properties)
Creator's jobtitle	Contains the job title of the photographer. As this is sort of a qualifier the Creator element has to be filled in as mandatory prerequisite for using Creator's Jobtitle		<a href="https://schema.org/creator-&gt;jobTitle">https://schema.org/creator-&gt;jobTitle</a>
Credit Line	The credit to person(s) and/or organisation(s) required by the supplier of the image to be used when published. This is a free-text field.		<a href="https://schema.org/creditText">https://schema.org/creditText</a> (in DRAFT)
Date Created	Designates the date and optionally the time the content of the image was created rather than the date of the creation of the digital representation.	Date and time of original data generation / DateTimeOriginal (36867, 0x9003) + OffsetTimeOriginal(36881, 0x9011)	<a href="https://schema.org/dateCreated">https://schema.org/dateCreated</a>
Description	A textual description, including captions, of the image.	if Image title / ImageDescription (270, 0x010E) contains a description and not a (short) title, or UserComment ( 37510, 0x9286)	<a href="https://schema.org/caption">https://schema.org/caption</a>

© 2020 IPTC (www.iptc.org) All rights reserved

25

25



## Helpers for getting out (2)

- Makers of software supporting to edit photo metadata** (for single computers, for a server-based system) **should consider**:
  - Keeping defined IPTC and Exif photo metadata values in sync (mostly done!!)
  - A function creating structured schema.org metadata for web pages.
- Makers of web content management systems (CMS)** **should consider**:
  - Importing IPTC and Exif metadata correctly to their database
  - A function transforming embedded IPTC and Exif metadata to structured schema.org metadata for a web page showing the photo.
  - Not to remove embedded metadata without any notice
  - Not to remove relevant fields like Creator, Copyright Notice, Credit Line, Description, Web Statement of Rights (Copyright URL), Licensor URL
- Makers of tools** (modules, libraries) **for web sites** **should consider** developing tools for an easy retrieval and display of embedded metadata.

Customers and users of this software: encourage the makers to consider this!

© 2020 IPTC (www.iptc.org) All rights reserved

26

26



## Helpers for getting out (3)

- To improve the **delivery speed** of photos on the web **is not a good reason for stripping off all metadata** (“slim files are delivered faster”)
- Make metadata of **photos on the web** more transparent = **display the fields**
  - ➔ Google Images makes them transparent in search results.
- CIPA is currently discussing with IPTC which **Exif Tags** set at the time of taking an image **shall be protected and conveyed endlessly**. Software makers will be encouraged to implement that.
- What about **making saved/embedded photo metadata secure** against not permitted changes?
  - What about a **history of metadata changes** which cannot be destroyed?
  - ➔ The Adobe Content Authority Initiative (CAI) has plans to achieve that.

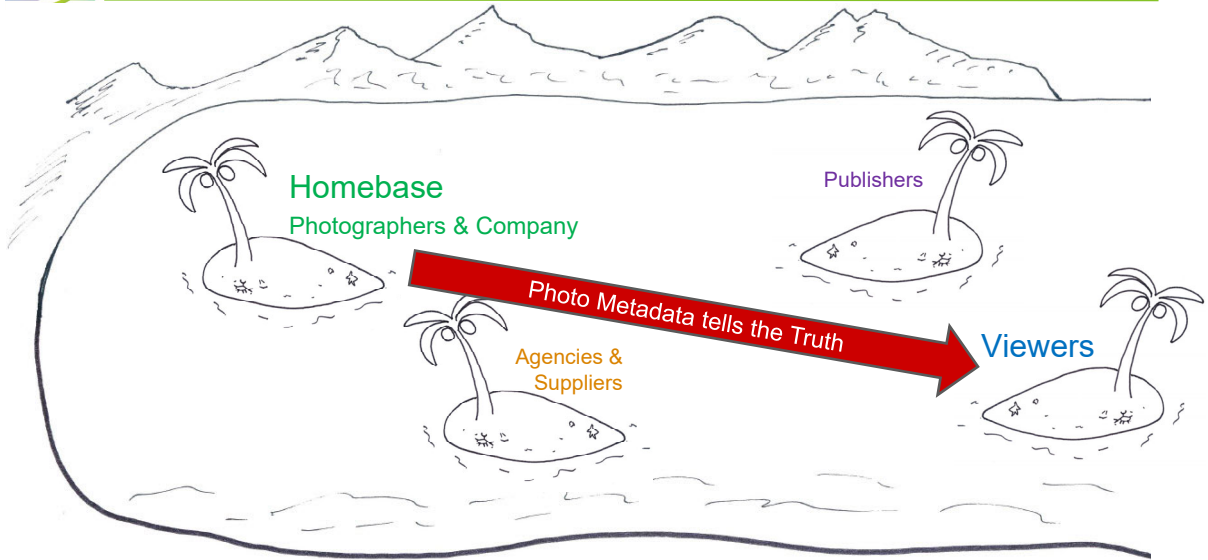


## Summary for photo businesses

- Take care to set correct values in the metadata fields
- Check if they are embedded correctly in your files
- Check if they are in photo files on any website controlled by you ... and IPTC tools help you checking this.
- If not: select a better software
- ... or talk to the maker of your software and request an improvement.
- Check if metadata gets lost somewhere in a supply chain you deliver to
- Remind customers not to remove metadata when publishing your photos
- Keep your eyes open for new things helping to keep metadata alive – and support this thing.



## Let's all be behind this goal



© 2020 IPTC (www.iptc.org) All rights reserved

29