



Workshop #5:
Working with Images in
Research

Agenda

Part I (Mathias Wyser, Samantha Foulger)

- How to organize your images
- Basics about Creative Commons Licences
- Re-using images for scientific publications

Part II (Roland Suri, Fabian Schneider)

- Let's dive into image formats

How to organize your images

Why manage images?

- It is easy to take and store a lot of pictures, and it is likewise easy to lose the overview.
- Images (not always just your own!) are a part of your research work. Therefore, it is best to organize them, the way you (should) organize all your other data as well.
- Image management saves time when you need to find a specific image, or if you need to know the rights when you have to publish an image.
- **While managing images is important, it is even more important to save them** – store additional copies of your work (images AND metadata).

Managing images: How?

Keep it simple!

1. Use a unique name/string for each image (=unique ID).
2. Add the necessary information (=metadata).
3. Use a simple system.

Simple Workflow for Image Management

Image Management

1. Create «good quality» images



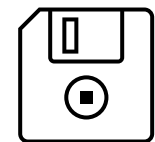
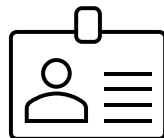
2. Rename picture with unique ID



3. Add metadata:
- Context
- Copyright info



4. Save both pictures and metadata!



Create «good quality» images - What is a «good quality» image?



The reason behind «good quality» is that the images should be in a format and size that can be used for publication and possibly long-term archiving.

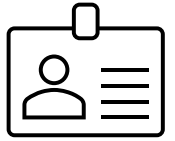
General recommendations:

- Format: TIFF / RAW / JPEG (high-quality)
- Color depth: 8 bit
- Resolution: 300 dpi
- No image compression



Quelle Foto: <https://www.saxoprint.de/blog/druckaufloesung-von-druckdaten>

2. Rename picture with a unique ID



- With a unique name (=record name), the metadata and your images match. This way you avoid duplicates.
- Use a consecutive number → sorting
→ **0000039 / 0000040 / ...**
- Use date and time
→ **202111170746** (yyyy/mm/dd/tttt)
- Use a simple system that helps you identify the main content/origin of the image, e.g. your initials before a number for your own images, images taken by others you mark with something else
→ **SF_15489** or **XX02936**
- Don't use spaces or special characters/symbols for your record names! → long term archiving

3. Add metadata: What is important?



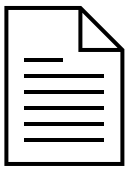
- **Metadata helps you to retrieve images! No metadata = no search results.**
- **Amount of metadata?**



The «right» metadata depends a lot on your field of studies and the use of the images.

- **Main information:** Record name, photographer / creator, keywords
- **Image content:** Short description/notes
 - Why and where was this image taken? Who/what is in it? → For publications: Did people in the image agree to a publication?
- **Technical metadata:** is usually saved automatically in the image information (date, time, lens type, GPS, etc.).

3. Add Metadata: What is important?



- **Image rights**

Be aware that most images are protected under Copyright Law.

If you manage images that are not your own, note who owns the rights. If you don't know, write down that the rights are unknown to you. This way, you won't reuse an image that you are not allowed to.

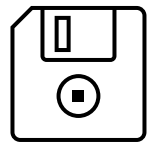


Publication: Was this image already published? If yes, who owns the rights?

- **License:** If the image is licensed, add this information.

- **Credits:** The credit is needed if you want to quote an image. Even if the image is not protected or free to use, always add this information.

4. Save both, pictures, and metadata! – How?



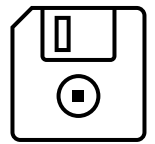
- **There are several tools to store your metadata and images**


- ➔ DAM software (professional)
- ➔ Image Management Software (private)
- ➔ Spreadsheet

Following some examples

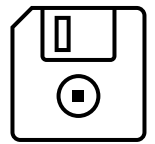


Example DAM Software



	Record Name: 0000000001590_0001
	Name: Pachycardien-Tuff
	Creator: Albert Heim
	Caption: Pachycardien-Tuff, submariner Tuff mit Pachycardia rugosa. Keuper. St. Cassianer und Raibler Schichten. Ladinian- [..]
	Location: Seiser Alpen (Alpi di Siusi)
	Period: Trias
	Categories: 03_Vulkanische Produkte
	Credits: ETH Zürich
	Publication Proof:
	License: CC BY-SA 4.0
	Copyright Notice: Erdwissenschaftliche Sammlung der ETH Zürich
	File Format: TIFF-Bild
	File Data Size: 19,806 KB
	DOI Link: http://doi.org/10.18748/ethz-a-000000866

Example Adobe Bridge



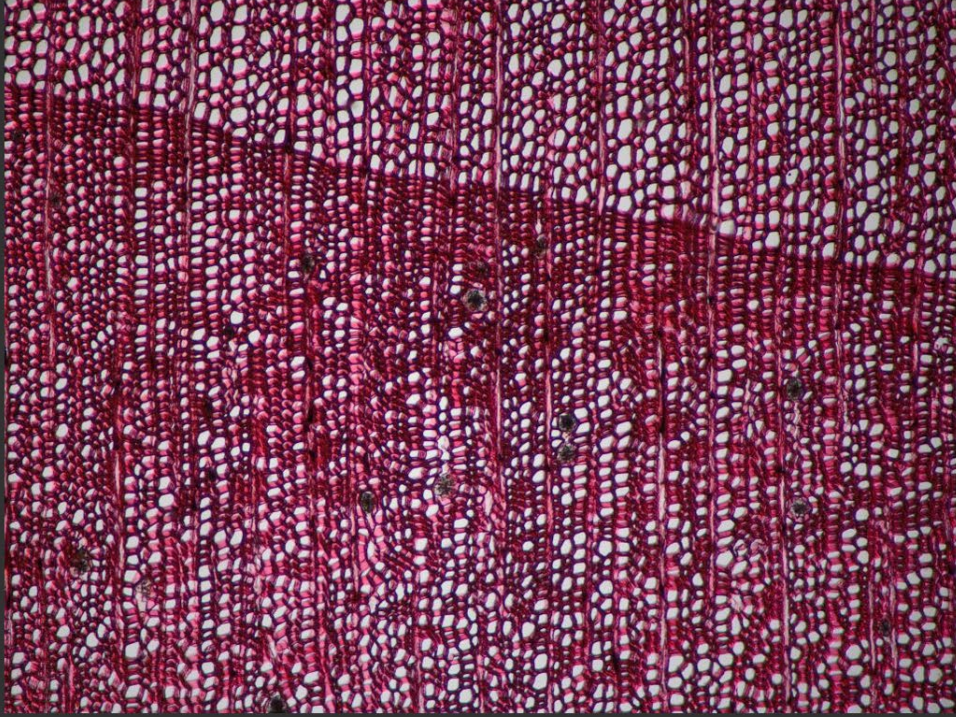
Desktop

Dieser PC

Metadaten

<i>f/</i> -- --	1024 x 768
-- --	325 KB 72 ppi
-- ISO --	Ohne Tags RGB

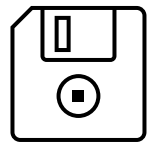
Ersteller	Institut für Baustoffe; Xylothek; ETH Zurich
Überschrift	Ginkgoaceae
Beschreibung	Querschnitt, Vergrößerung 40 x, Jahrringgrenze ziemlich deutlich. Zwischen den Wanddicken der Früh- und Spätholztracheiden besteht kaum ein Unterschied. Runde Kristalldrüsen
Stichwörter	Ginkgo biloba L.
IPTC-Subject Code	
Erstellungsdatum	
Gattung	
IPTC Scene Code	
Titel	
Credit	
Quelle	
Copyright-Vermerk	Xylothek, Institut für Baustoffe (IfB) der ETH Zürich
Copyright-Status	Durch Copyright geschützt
Nutzungsbedingungen	CC BY-SA 4.0



Gymno_0264_01.jpg

Inhalt

Example Spreadsheet



	A	B	C	D	E	F	G	H	I	D
1	Bildcode	Titel	Material	Objektmasse	Gewicht	Beschreibung	Inschrift	Zustand	Datum 1	
2	ETHZ_IFHT_0001	Messbrücke W 51 8605-6	Metall (allgemein)	30.5 x 18.5 x	5326 g	Die Wheatstonesche Messbrücke	gut		1913	
3	ETHZ_IFHT_0001_a	Messbrücke W 51 8605-6								
4	ETHZ_IFHT_0001_b	Messbrücke W 51 8605-6								
5	ETHZ_IFHT_0001_c	Messbrücke W 51 8605-6								
6	ETHZ_IFHT_0002	Messbrücke W 51 8605-1	Metall (allgemein)	30.5 x 18.5 x	5261 g	Die Wheatstonesche Messbrücke	gut		1913	
7	ETHZ_IFHT_0002_a	Messbrücke W 51 8605-1								
8	ETHZ_IFHT_0002_b	Messbrücke W 51 8605-1								
9	ETHZ_IFHT_0002_c	Messbrücke W 51 8605-1								
10	ETHZ_IFHT_0003	Messbrücke 14 11d	Metall (allgemein)	32.5 x 14.5 x	3454 g	Die Wheatstonesche Messbrücke	gut		1880	
11	ETHZ_IFHT_0003_a	Messbrücke 14 11d								
12	ETHZ_IFHT_0003_b	Messbrücke 14 11d								
13	ETHZ_IFHT_0003_c	Messbrücke 14 11d								
14	ETHZ_IFHT_0004	Messbrücke L. E.	Metall (allgemein)	33 x 16 x 16 c	3647 g	Die Wheatstonesche Messbrücke	gut		1880	

Discussion: What did you do so far with your metadata? What are you (probably) going to change?

Consistency of Metadata

- Make sure to always use the same order/rule when entering metadata:

- Miller, Henry
- Henry Miller
- Miller, Henry (Photographer)

Old system	New System	New System
Name	First name	Family name
Miller, Henry	Henry	Miller
John P., Doe	Doe	John P.
Jane Doe		Jane Doe

- Consistency in your metadata makes it easy to filter the data and for other systems to read it (or change it), e.g. when importing your data.

Consistency of Metadata: Example

	A	B	C	D	E	F	G	H
1	Record name	Genus	Species	Author	Country	Date	Licence	
2	0058920	Melampsorium	hiratsukanum	S. Ito ex Hirats. f.	France	6-Jan-1916	CC BY SA	
3	0058958	Puccinia	adoxae	R. Hedw.	Switzerland	6-Jan-1916	CC BY SA	
4	0058959	Puccinia	aegopodii	(Schumach.) Link	Switzerland	29-Mar-1897	CC BY SA	
5	0058957	Puccinia	adoxae	R. Hedw.	Switzerland	01 Jan 1957	CC-BY-SA	
6	058956	Puccinia	singularis	Magnus	Switzerland	08 1966	CC-BY-SA	
7	0058954	Tranzschelia	fusca	(Wallr.) Dietel	Switzerland	8-07-1926	CC BY SA	
8	058938	Puccinia	crepidis-leontodontoidis	Maire	France (Corsica)	08-May-1977	Creative Commons BY SA	
9	0058939	Puccinia	crepidis-leontodontoidis	Maire	France (Corsica)	8-May-1978	CC BY SA	
10								

The «Masterfile»



- Store **one masterfile** with a good resolution (if possible **Tiff**).
- If you need a smaller image later, you can convert the masterfile.
- If you want/need to store different sizes/formats of the same image, name them in a way so you can recognize immediately if it is a masterfile or not (e.g. ND_873_Master, ND_873_changed20201001).

Use a simple system

- Use a system that is easy to use, that you are already familiar with or use anyway.
- A simple spreadsheet is good and usually easy to import into another systems if necessary.

Image ID	Title	Photographer	Description	Rights	Keywords
ND_5543_9.Tiff	Diagram 31	Doe, John P.	Assembly diagram: the proposed modular system can be adapted for custom slab dimensions and post-tensioned on both directions	Doe, Jane; Institut XYZ; ETH Zurich	diagrams; Institut XYZ; Jane Doe;

- Adobe Bridge is a product specifically designed for image management and it is free.

Metadata: In the image file or separately

- Using a spreadsheet: Metadata is completely separate from the images. No limit to the amount and type of metadata fields.
- Using Photoshop/Bridge/Lightroom: Metadata is stored IN the image (IPTC). A given amount of metadata fields which can be chosen from.
- Advantages of storing metadata within the images:
 - No separate file/storage is needed, all is in the same place.
- Disadvantages of having the metadata in the images:
 - Export of metadata in Adobe products does only work with an extra script.

Adobe Bridge

- Useful for basic image management (<https://www.adobe.com/products/bridge.html>)
- The images stay on your hard drive, in the folders you already stored them in.
- Metadata is saved IN the image, not separately.
- Core. For more information on IPTC: <https://www.iptc.org/std/photometadata/specification/IPTC-PhotoMetadata>
- Be aware: Limited number of characters! <https://de.wikipedia.org/wiki/IPTC-IIM-Standard>
By-line (author): 32 / Credit: 32 / Copyright notice: 128 / Keywords (each): 64
- A lot of information and videos about Bridge can be found on the internet.

E-Pics

- Image Management for ETH entities (e.g. research groups – not for personal use), free or charge.
- Based on a digital asset management system (DAM)
- Possibility to share your images via backend or via frontend (for larger groups)
- Image catalogue is set up individually according to the customers needs (i.e. metadata fields)
- www.e-pics.ethz.ch

Break

Basics about Creative Commons Licences





Why use Creative Commons licences?

- **If you do not use a licence, your work is automatically copyright protected**
= no one else can use it without your permission.
- Creative Commons licences make it easy to allow others to reuse your images.
- The licences are clearly defined and easy to understand.
- To use the licences, you **must** be the copyright holder!
- Once you publish an image with a CC licence, you cannot withdraw the licence.
- <https://creativecommons.org/>

Creative Commons licences

- The licences consist of the following parts:

- Always  (=stands for Creative Commons) with one or several of the following signs:

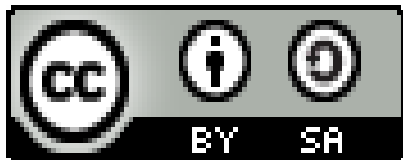
BY 	By	Credit must be given to the creator.
SA 	Share Alike	Adaptions must be shared under the same terms.
NC 	Non Commercial	Only noncommercial uses of the work are permitted.
ND 	Non Derivative	No derivatives or adaptations of the work are permitted

Creative Commons Licences: Definitions

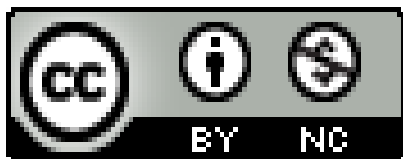
(<https://creativecommons.org/>)



CC BY: This license allows reusers to **distribute, remix, adapt, and build upon** the material in any medium or format, so long as **attribution is given to the creator**. The license allows **for commercial use**.



CC BY SA: This license allows reusers to **distribute, remix, adapt, and build upon** the material in any medium or format, so long as **attribution is given to the creator**. The license allows for **commercial use**. If you remix, adapt, or build upon the material, you must **license the modified material under identical terms**.

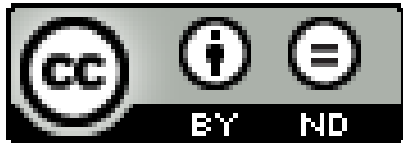


CC BY NC: This license allows reusers to **distribute, remix, adapt, and build upon** the material in any medium or format for **noncommercial purposes only**, and only so long as **attribution is given to the creator**.

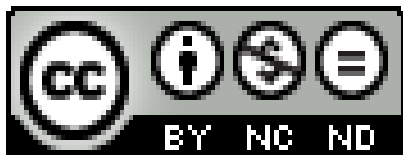
Creative Commons Licences: Definition (by CC)



CC BY-NC-SA: This license allows reusers to **distribute, remix, adapt, and build upon** the material in any medium or format for **noncommercial purposes** only, and only so long as attribution is given to the creator. If you remix, adapt, or build upon the material, you must **license the modified material under identical terms**.



CC BY-ND: This license allows reusers to copy and distribute the material in any medium or format in **unadapted form** only, and only so long as **attribution is given to the creator**. The license allows for commercial use.



CC BY-NC-ND: This license allows reusers to copy and distribute the material in any medium or format in **unadapted form** only, for **noncommercial purposes** only, and only so long as **attribution is given to the creator**.

Creative Commons licences: Public Domain - Differences



CC Zero: With this licence the copyright holders **gives their work into the public domain**, i.e. they waive the copyright on this specific work **completely**. Reusers can distribute, remix, adapt, and build upon the material.



Public Domain Mark: This licence states that the work is no longer protected by copyright (e.g. due to its age).

Creative Commons licence: What to consider

If you want to licence your own image, be aware of the following:

- You need to be copyright holder.
- You cannot revoke the CC licence.
- Use a link to the licence you chose.
- Always state the licence with its version (CC BY-SA 4.0 vs. CC BY-SA 2.0).
- Follow best practices for attribution.
- Be careful with the NC licence.

Creative Commons licence: Correct attribution

Image Information

Record Name: Dia_009-091
Photographer: Heim, Arnold
Title: Zoo, Sydney
Caption: Reisen in Australien und Tasmanien, 1920-21. Känguruh
Dating: 29.11.1921
Physical Description: Photography : glass-plate positive
Colour: hand coloured
Orientation: Horizontal
Format: 8,5 x 10 cm
Categories: Heim, Arnold Heim, Arnold (1882-1965, geologist) Kangaroos + Koalas Nature photography Sydney
Keywords: beutel beuteltier känguru Natur niedlich Säugetier Tier Tierwelt wallaby wild zoo
File Data Size: 24,586 KB

Terms of Use

License: CC BY-SA 4.0
Availability: Free download and use
Copyright Notice: ETH-Bibliothek Zürich, Bildarchiv / Fotograf: Heim, Arnold / Dia_009-091 / CC BY-SA 4.0
DOI Link: <http://doi.org/10.3932/ethz-a-000025410>

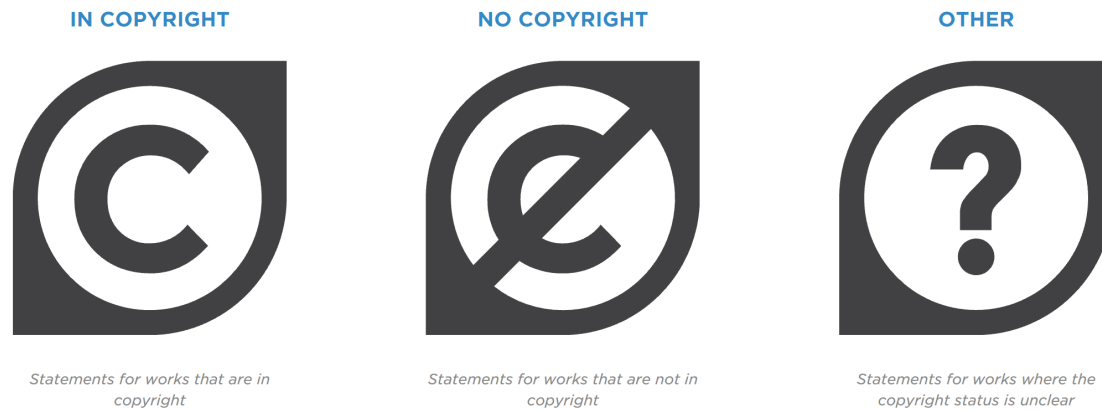
Title	Zoo, Sidney
Author	Heim, Arnold
Source (plus Link to source)	ETH-Bibliothek Zürich, Bildarchiv http://doi.org/10.3932/ethz-a-000025410
Licence (plus link to licence)	CC BY-SA 4.0

- https://wiki.creativecommons.org/wiki/Best_practices_for_attribution

Discussion in groups: Which Creative Commons Licences could/would you use for your images?
Why?

Other Licences: RightsStatements.org I

- <https://rightsstatements.org/en/>
- Similar to Creative Commons Licences, but mostly for cultural heritage



- Where Creative Commons licences are not enough or do not fit, maybe the licences from Rights Statements do.

Other Licences: RightsStatements.org II

- The licences are always a combination of the three categories plus text.
- In copyright, but rights-holder(s) unlocatable or unidentifiable.



- In copyright, but rights-holder(s) make the items available for educational purposes.



Other Licences: RightsStatements.org III

- Public domain, but cannot be re-used freely. Important: Always check in detail with the organization who provides this item what this means in detail!



- Unclear situation, copyright status is unknown but was also not researched. Always check in detail with the organization who provides this item what this means in detail!



(Re-) Using images for scientific publications

How to re-use published images

- If you want to re-use an already published image in a future publication, and there is no (Creative Commons) licence, you have to ask the publisher for permission – **even if it is an image you have taken yourself**. Usually, the rights are with the publisher.
- Look for RightsLink Services: <https://pubs.acs.org/doi/10.1021/acsearthspacechem.9b00223#>

RETURN TO ISSUE | < PREV ARTICLE NEXT >

Processes Governing Chromium Contamination of Groundwater and Soil from a Chromium Waste Source

Mainak Bhattacharya, Amritanshu Shrivastav, Shrikant Bhole, Rahul Silori, Tim Mansfeldt, Ruben Kretzschmar, and Abhas Singh*

✓ Cite this: *ACS Earth Space Chem.* 2020, 4, 1, 35–49

Publication Date: December 13, 2019 ▾

<https://doi.org/10.1021/acsearthspacechem.9b00223>

Copyright © 2019 American Chemical Society

[RIGHTS & PERMISSIONS](#) ✓ Subscribed

PDF (8 MB)

open URL

Article Views

430

Altmetric

1

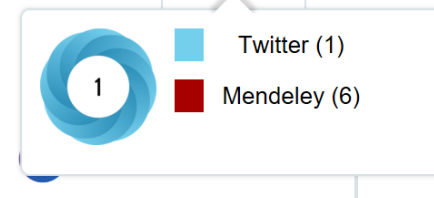
Citations

2

Share

Add to

Export

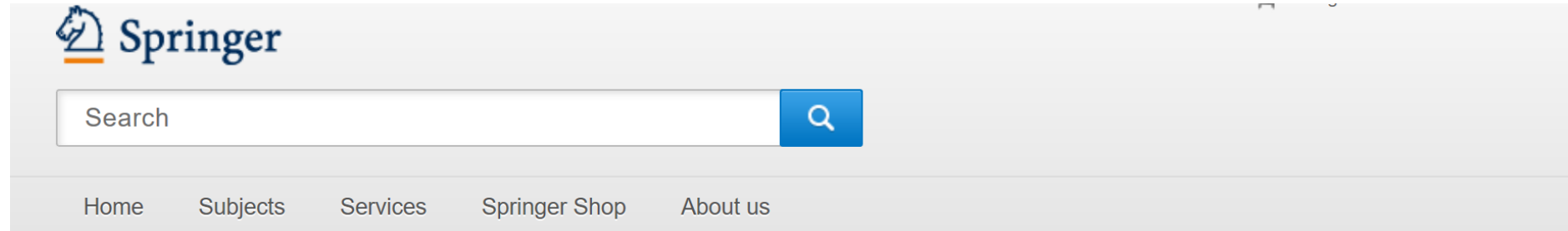


SUBJECTS: Chromium, Soils, ▾

How to re-use published images

- If there is no link, contact the publisher in another way. It is very important to get written permission if you want to re-use an already published image! Every publisher is different, but information on how to proceed can usually be found.

Example:



Rights and Permissions

- » [Permissions](#)
- » [Obtaining translation and reprint rights](#)
- » [Rights and Permissions at book fairs](#)
- » [Contacts](#)
- » [Anti-piracy strategies for Springer eBooks](#)
- » [Springer's text- and data-mining policy](#)

Permissions

Get permission to reuse Springer Nature content

Springer Nature is partnered with the Copyright Clearance Center to meet our customers' licensing and permissions needs.

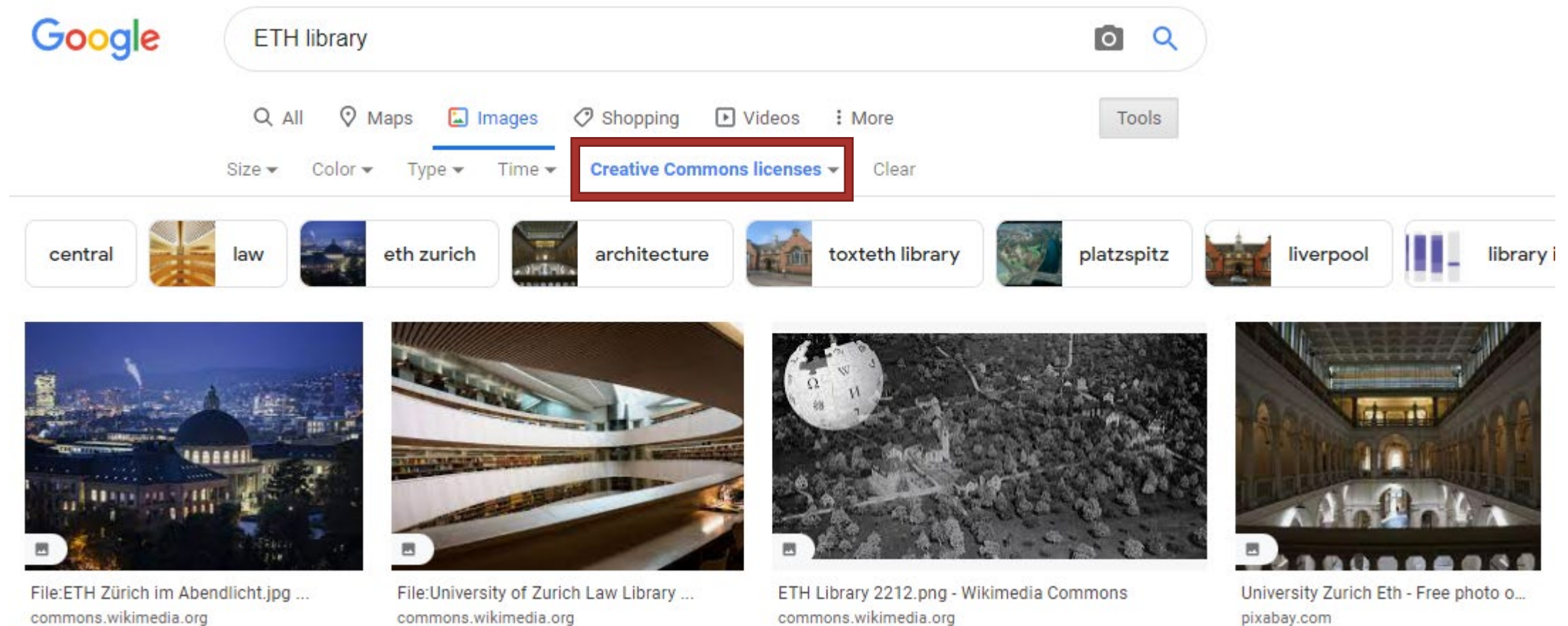
Copyright Clearance Center's RightsLink® service makes it faster and easier to secure permission for the reuse of Springer Nature content to be published, for example, in a journal/magazine, book/textbook, coursepack, thesis/dissertation, annual report, newspaper, training materials, presentation/slide kit, promotional material, etc.

Simply visit [SpringerLink](#) and locate the desired content;

- Go to the article or chapter page you wish to reuse content from. (Note: permissions are granted on the article or chapter level, not on the book or journal level). Scroll to the bottom of the page, or locate via the side bar, the "Reprints and Permissions" link at the end of the chapter or article.
- Select the way you would like to reuse the content;

Where to find images with licenses: Google

- Google Search: Filter by licence (via «tools»)

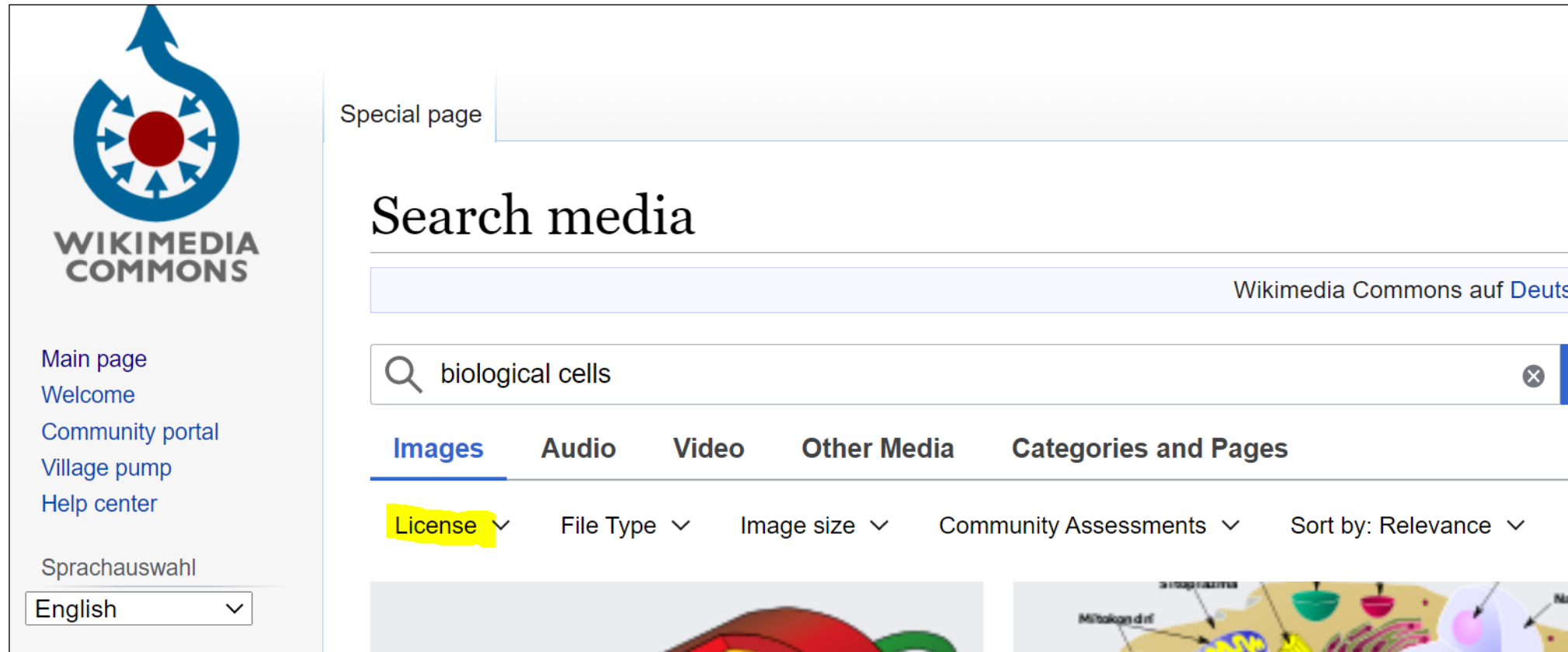


The screenshot shows a Google search for "ETH library" on the Images tab. The "Creative Commons licenses" filter is highlighted with a red box. Below the search bar, there are several image thumbnails with their respective captions and sources:

- central**: File:ETH Zürich im Abendlicht.jpg ... commons.wikimedia.org
- law**: File:University of Zurich Law Library ... commons.wikimedia.org
- eth zurich**: ETH Library 2212.png - Wikimedia Commons commons.wikimedia.org
- architecture**: University Zurich Eth - Free photo o... pixabay.com
- toxteth library**
- platzspitz**
- liverpool**
- library i**

Where to find images with licenses: Wikimedia Commons

- <https://commons.wikimedia.org/>



The screenshot shows the Wikimedia Commons search interface. On the left is a navigation sidebar with the Commons logo and links to Main page, Welcome, Community portal, Village pump, Help center, and Sprachauswahl (English). The main content area is titled 'Search media' and features a search bar with the query 'biological cells'. Below the search bar are tabs for 'Images', 'Audio', 'Video', 'Other Media', and 'Categories and Pages'. Under the 'Images' tab, there are filter options: 'License' (highlighted in yellow), 'File Type', 'Image size', 'Community Assessments', and 'Sort by: Relevance'. The bottom of the page shows the beginning of search results, including a diagram of a cell with labels like 'Mitochondrien' and 'Chloroplasten'.

Where to find images with licenses: Pixabay

- <https://pixabay.com/de/>



Where to find images with licenses: Tipps

Important: To be sure about the licence / conditions of reuse, **always check the original source** and make sure the licence is the same!

Re-use the image exactly in the way you are allowed to, and add the credits the publisher tells you.

Image management: Mention in your metadata which of your images have been published and where.

Samantha Foulger, Mathias Wyser
e-pics@library.ethz.ch

ETH Library
E-Publishing
HG H 31.5
Rämistrasse 101
8092 Zurich

www.e-pics.ethz.ch



Workshop #5: Working with Images in Research

Let's dive into image formats (Roland Suri & Fabian Schneider)

Agenda

1. How to organize your images (Samantha Foulger)
2. Basics about Creative Commons Licences (Samantha Foulger)
3. Re-using images for scientific publications (Samantha Foulger)
4. Let's dive into image formats (Roland Suri & Fabian Schneider)



Agenda

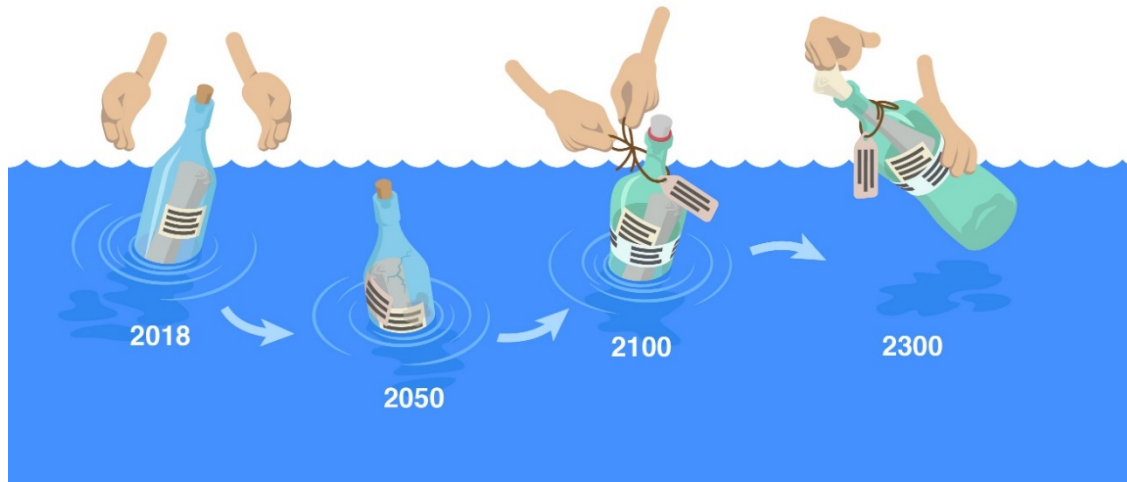
- Image Formats
- TIFF (Tagged Image File Format)
- Defect Files

Image Formats

Fabian Schneider

Significant Properties and Designated Community

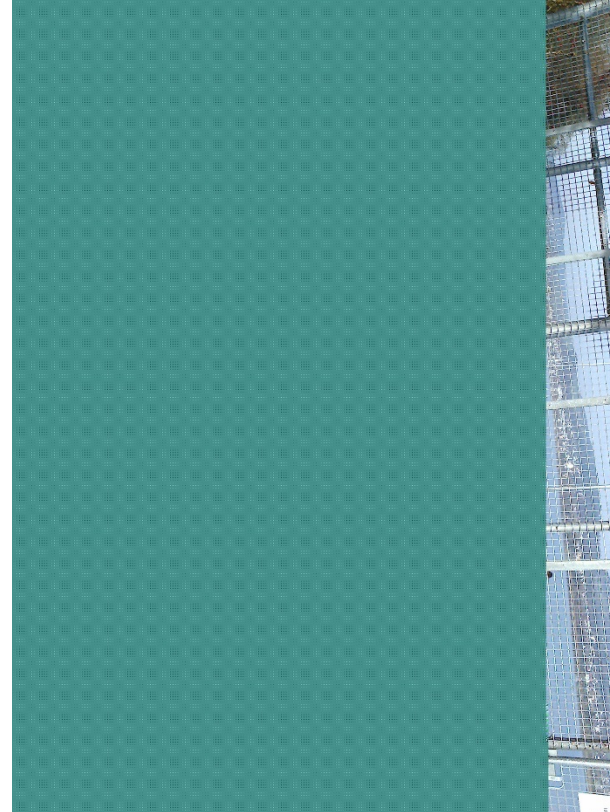
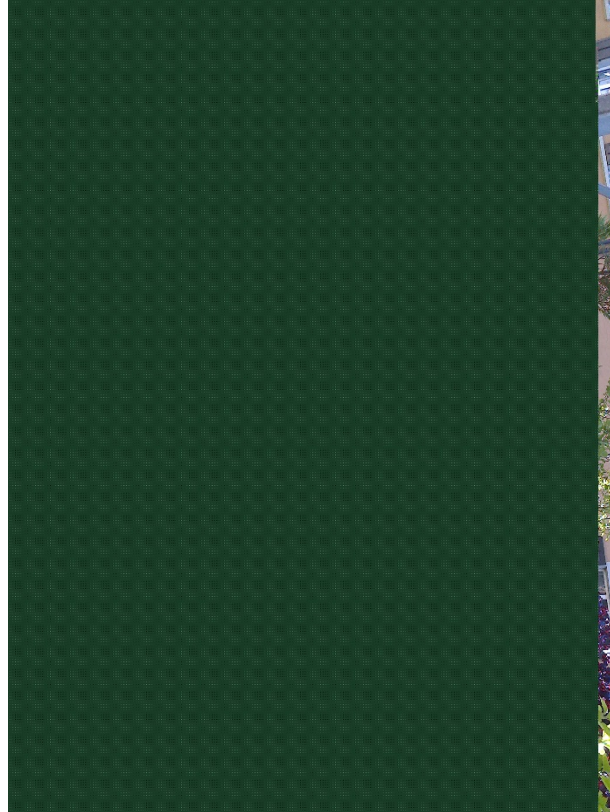
- In Digital Preservation we want to:
 - Safeguard data
 - Ensure accessibility and reusability
- The «**designated community**»: Who will reuse my data?
- The «**significant properties**»: What is needed to reuse my data?



Why should we think of digital preservation?

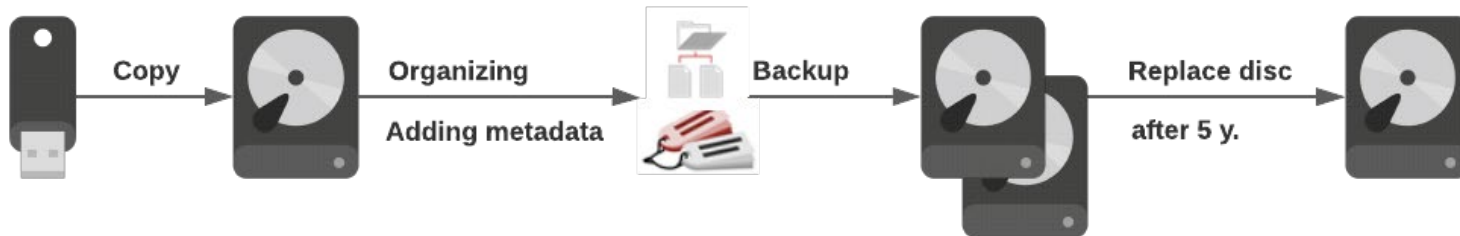
Example

- Found a flash drive with these pictures on it:



Real Life Example Hardware Defect

- What could have been done to prevent this data loss?
 - I should have **copied** the pictures on a harddrive!
 - If I had **organized** these pictures, I wouldn't have forgot them!
 - I should have done **backups**!
 - As every drive has a **limited lifetime**, i should replace the disc!



- Lifetime of media:

Magnetic Harddisk	5 yrs. (error probability increases)
USB flashdrive	10 yrs.
CD/DVD	30 yrs.

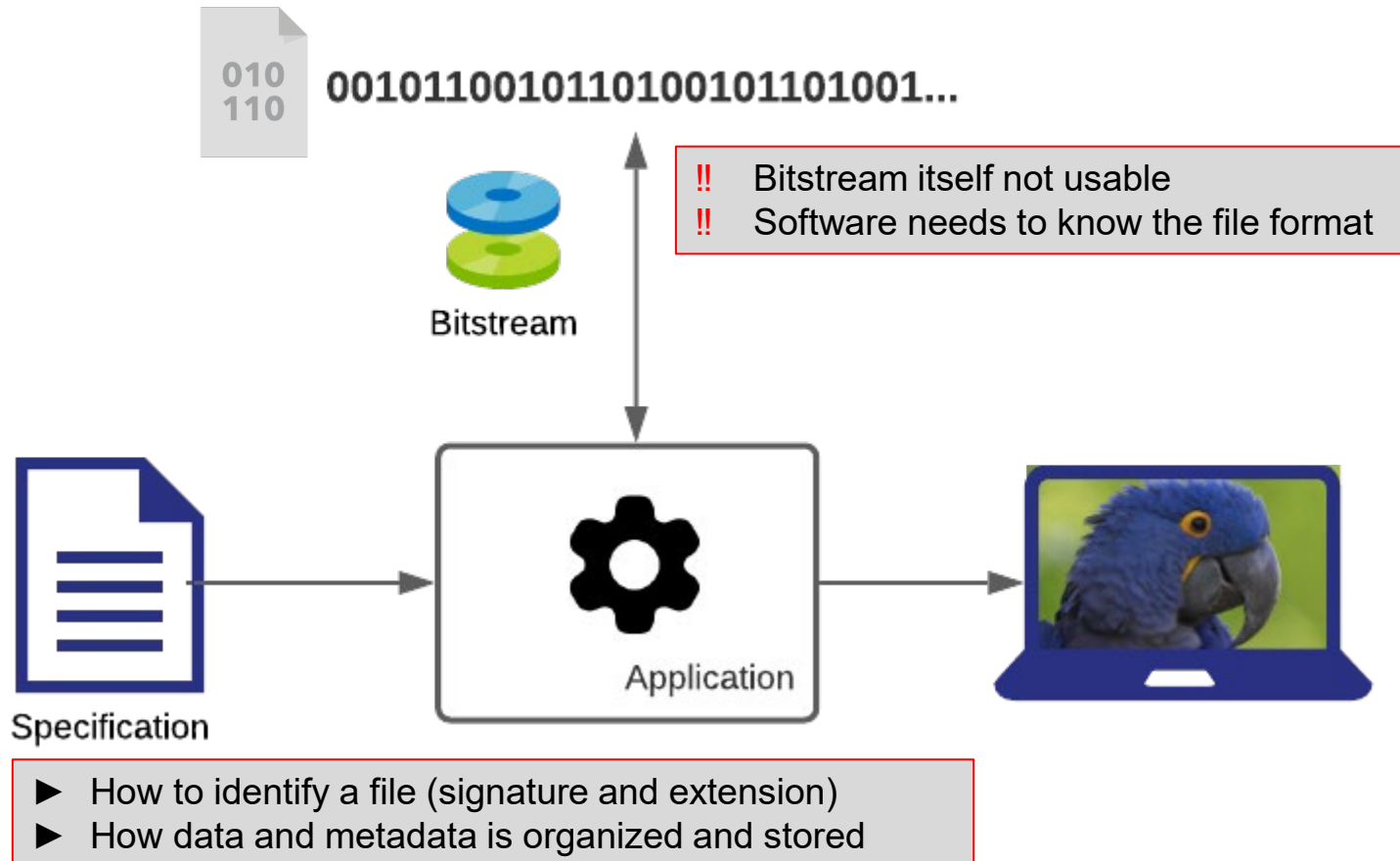
Risks in Digital Preservation

- Risks are impacts or conditions which could prevent us from reusing data in two ways:
 - Data is technically well, but not usable for some reasons*
 - Data is irreversible corrupt and cannot be accessed anymore**

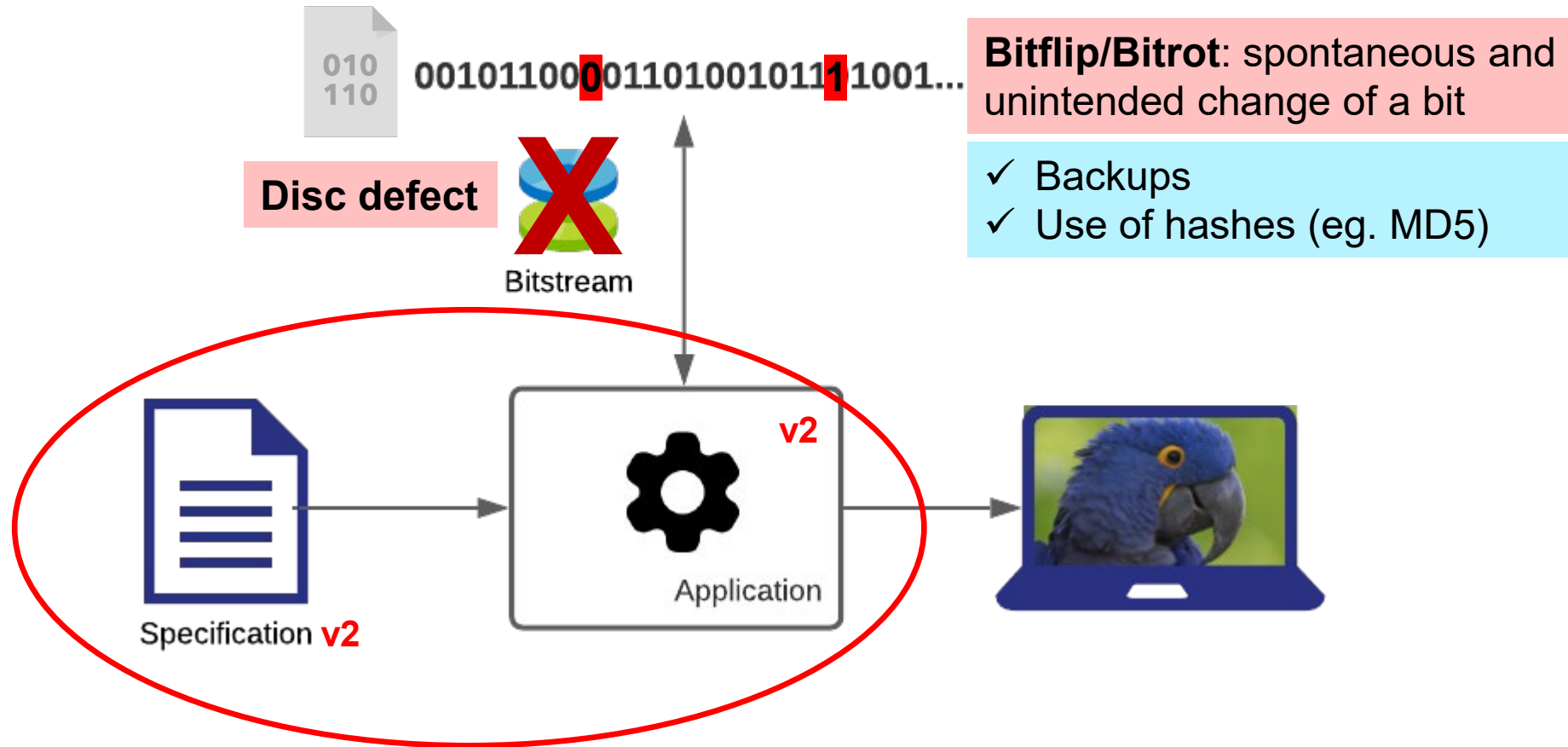
Risk	Mitigation
Data deluge*	Keep organized
Missing context information*	Provide metadata
Hardware failures**	Backups and integrity checks
Obsolescence*	Keep track on file formats you have and which one you choose



Bitstreams and Formats



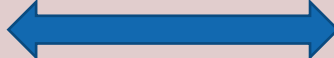

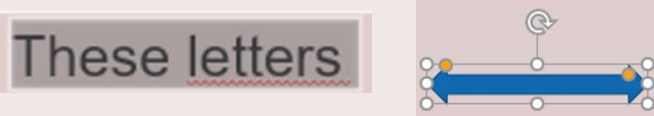


Bitstreams and Formats



New file format or software version

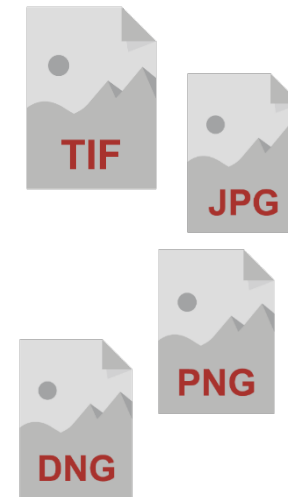
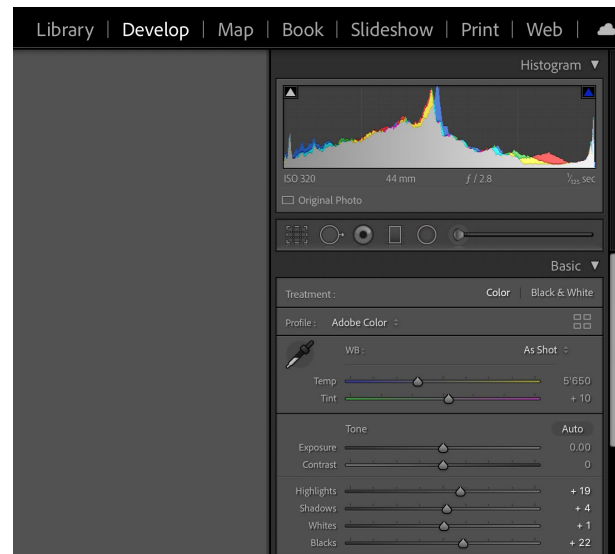
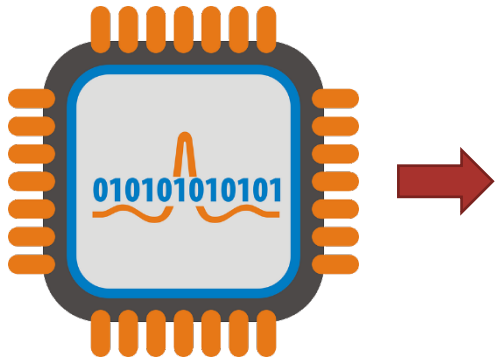
→ Keep track of your data and convert files if necessary

Two Types of Images

	Vector	Raster/Bitmap
Example	<p>These letters, fonts in general Shapes like this:</p> 	
Characteristics	<ul style="list-style-type: none"> • Small filesize • Lossless enlargement • Images can be selected: 	<ul style="list-style-type: none"> • Large filesize (compared to vector) • Loss of resolution when enlarged • Objects cannot be selected (except some fancy AI)
Formats e.g.	SVG 	TIFF, JPG, PNG, BMP 

RAW data

- RAW itself is not a specific format and (except DNG) proprietary
- DNG (Digital Negative, first developed 2004 by Adobe) is an open format
- RAW-data comes directly from the sensor and needs to be processed
- Specific RAW converters for each RAW format or Adobe Camera Raw
- The result should be saved in another format, preferably TIFF



RAW-data e.g. DNG, NEF, CR2 etc.

RAW vs. DNG

RAW formats

- Proprietary
- Not writeable
- Metadata could be used by manufacturer's RAW converter
- Sometimes no thumbnails in OS



DNG

- Open format
- Metadata can be saved within DNG
- Probably loss of manufacturer's specific metadata
- Better compatibility because of open specification



General Formats Recommendations

- Are open source / specification is available
 - Specification allows understanding the format
 - Specification allows validating a format
- Are widespread
 - Longer support
 - More tools at your disposal
- Provide lossless compression (better no compression at all)
 - Less complexity
 - Best quality (for reuse and migration)



Image Formats Recommendations

Recommended	<ul style="list-style-type: none">• TIFF (8 bit) (most important for preservation)• DNG, TIFF (16 bit) (for preserving raw data/editing)• JPG2000 (lossless) (sharing only)• PNG (lossless)• SVG (for vector graphics)
Acceptable	<ul style="list-style-type: none">• JPG• JPG2000, PNG (lossy compressed)• TIFF (with LZW compression)• GIF• BMP

► Whenever possible use TIFF Baseline, 8 bit

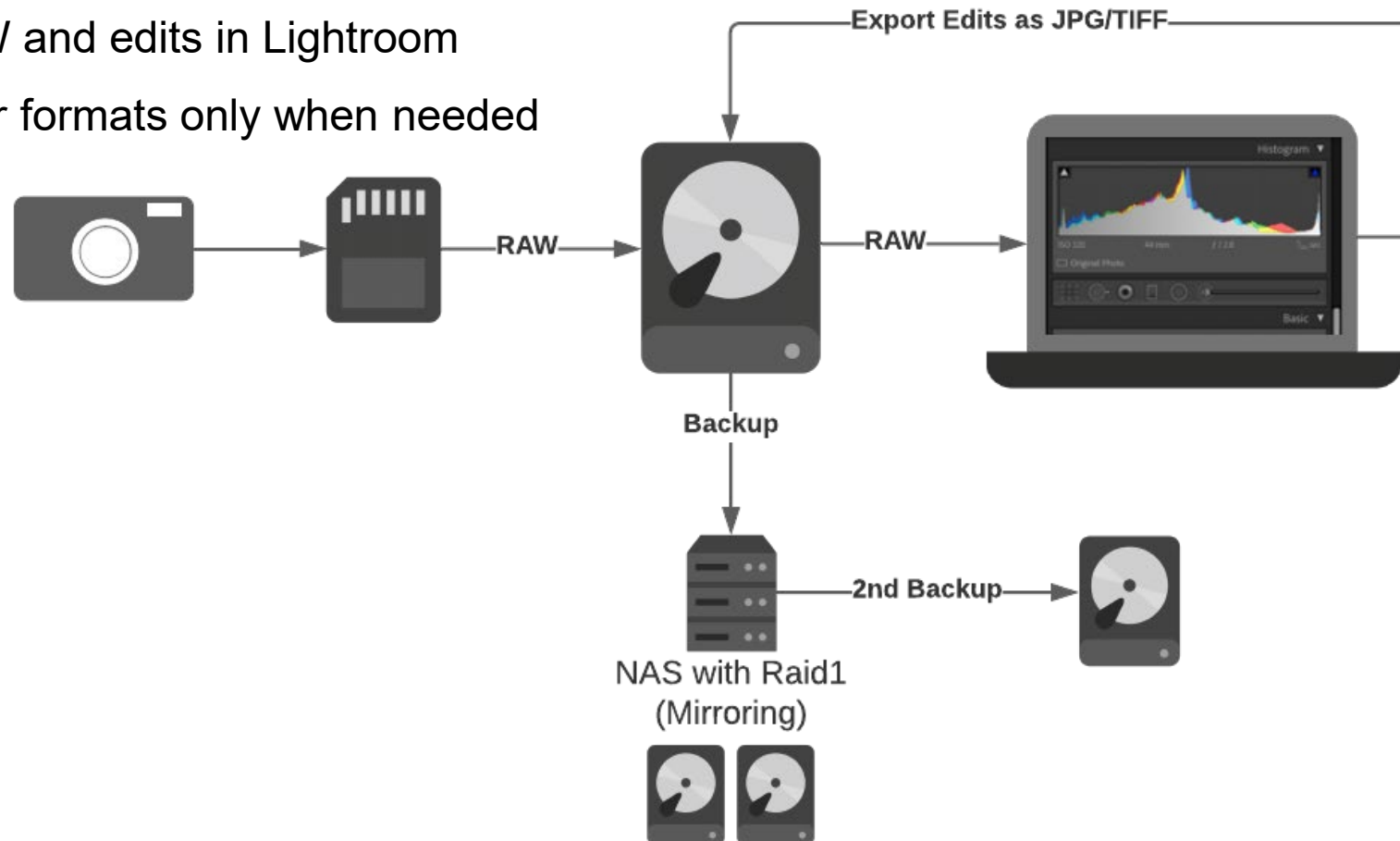
Recommended image file formats for long-term archiving

File type	Recommended	Suitable to only a limited extent	Not suitable for archiving
Raster image (bitmap)	<ul style="list-style-type: none">• TIFF (*.tif) (uncompressed, preferentially TIFF 6.0, Part 1: baseline TIFF). TIFF is preferred as compared to PNG or JPEG2000.• Portable Network Graphics (*.png, uncompressed)• JPEG2000 (*.jp2, lossless compression)• Digital-Negative-Format (*.dng) to keep raw data of digital fotos in addition to an second copy in TIFF format	<ul style="list-style-type: none">• TIFF (*.tif) (compressed)• GIF (*.gif)• BMP (*.bmp)• JPEG/JFIF (*.jpg)• JPEG2000 (lossy compression) (*.jp2)	
Vector graphics	<ul style="list-style-type: none">• SVG without JavaScript binding (*.svg)		<ul style="list-style-type: none">• Graphics InDesign (*.indd), Illustrator (*.ait)• Encapsulated Postscript (*.eps)• Photoshop (*.psd)

ETH Data Archive recommendations see <https://documentation.library.ethz.ch/display/DD/File+formats+for+archiving>

Example Workflow with Images

- As a photographer I need the most out of my images for postprocessing → RAW
- Different RAW-formats from different Cameras
- Keep the RAW and edits in Lightroom
- Export in other formats only when needed



Example Workflow with Images

- Keep RAW or convert to DNG/TIFF?
- I keep the original RAW because:
 - Productive system: I want to re-edit old images
 - Additional effort for converting to DNG
 - More storage required when choosing TIFF
- Therefore I keep track of my data:
 - Periodically check old RAW-files
 - Especially when RAW-Converter gets Updates
- In another use case, when edits have been finished I would save the results in TIFF 8 bit for preservation and probably de RAW data as DNG.

Format	Filesize
ARW (original RAW)	23.8 MB
DNG	25.4 MB
DNG lossy compressed	8.96 MB
DNG with RAW embedded	48.8 MB
DNG with RAW embedded, lossy compressed	32.3 MB
TIFF 8 bit, uncompressed	68.6 MB
TIFF 16 bit, uncompressed	137 MB

Example: My Workflow with Images

What I have in my Archive:

File Type	
All (8 File Types)	
Digital Negative / Lossless	DNGs without compression or embedded RAW.
Digital Negative / Lossy Compressed	Unfortunately I lost quality here for no reason! ☹️
JPEG	From my mobile phones, JPG has excellent support, no action required so far.
Photoshop Document (PSD)	Edits in Photoshop, I saved the results as TIFFs but kept the PSD!
PNG	I prefer PNG over JPG because of the support for transparency.
Raw	All my RAWs, not suitable for preservation, but periodically checked!
TIFF	The results of my PSDs, without compression or with LZW.

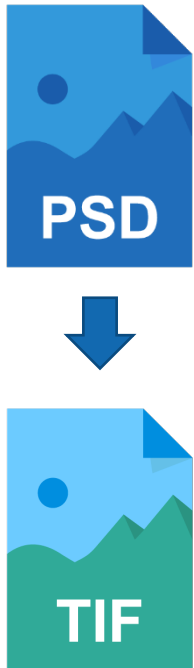
TIFF (Tagged Image File Format)

Roland Suri

Why TIFF for archiving bitmaps?

- Popular, well-established, non-proprietary
- Baseline TIFF (=TIFF 6.0, Part 1) as defined by Adobe Systems in 1992

Saving TIFF Images



TIFF Options

Image Compression

- None
- LZW
- ZIP
- JPEG

Quality: Maximum
small file large file

Pixel Order

- Interleaved (RGBRGB)
- Per Channel (RRGGBB)

Byte Order

- IBM PC
- Macintosh

Layer Compression

- RLE (faster saves, bigger files)
- ZIP (slower saves, smaller files)
- Discard Layers and Save a Copy

Save Image Pyramid

Save Transparency

Speichern: Als Kopie

- Anmerkungen
- Alpha-Kanäle
- Volltonfarben
- Ebenen

Farbe: Proof-Einstellung: CMYK-Arbeitsfarbraum

ICC-Profil: sRGB IEC61966-2.1

Andere: Miniatur

OK Cancel

About compression

- Different compression methods with different features
 - The more you compress, the longer you wait
 - Choose lossless if possible (LZW is widely used for TIFF)
 - Lossy only if there is no choice (JPG)
 - Do a visual check if you compress files
- Best is to avoid any compression !



Saving files

- Short names for files and directories (< 256 characters in Windows)
- In file and folder names no symbols: \ / ? : * " > < | : # % " { } | ^ [] ` ~ blanks

How to check the signature: DROID

DROID is a free app from nationalarchives.gov.uk

The screenshot shows the DROID v6.4 application interface. The main window displays a table of file signatures. The table has the following columns: Resource, Extension, Size, Last modified, Ids, Format, Version, Mime type, PUID, and Method. The table contains three rows of data:

Resource	Extension	Size	Last modified	Ids	Format	Version	Mime type	PUID	Method
C:\C temp\Bilder			23.11.20 10:44						
TIFF File old.tif	tif	1.4 MB	16.11.20 18:00		Tagged Image File Format		image/tiff	fmt/353	Signature
JPEG File.jpg	jpg	855.5 KB	22.08.15 19:58		Exchangeable Image File Format (Compressed)	2.2	image/jpeg	x-fmt/391	Signature
TIFF File without extension		1.4 MB	16.11.20 18:00		Tagged Image File Format		image/tiff	fmt/353	Signature

Annotations in the image:

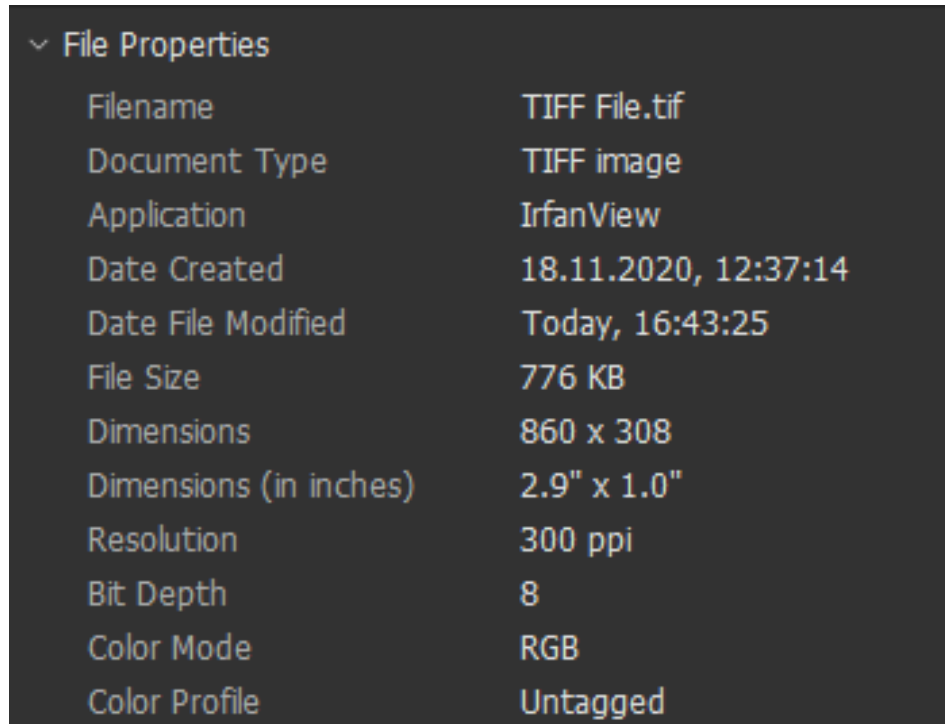
- 'File format' points to the 'Format' column.
- 'Link to format description at National Archives' points to the 'PUID' column.
- 'Identification used signature' points to the 'Method' column.

See also <https://documentation.library.ethz.ch/display/DD/File+formats+for+archiving>

Analyse single files: www.itforarchivists.com/siegfried

Technical Metadata in Adobe Bridge

- Image editors (Bridge, Photoshop) list the technical metadata for you

A screenshot of the 'File Properties' panel in Adobe Bridge. The panel is dark grey with white text. It lists various technical metadata for a file named 'TIFF File.tif'.

File Properties	
Filename	TIFF File.tif
Document Type	TIFF image
Application	IrfanView
Date Created	18.11.2020, 12:37:14
Date File Modified	Today, 16:43:25
File Size	776 KB
Dimensions	860 x 308
Dimensions (in inches)	2.9" x 1.0"
Resolution	300 ppi
Bit Depth	8
Color Mode	RGB
Color Profile	Untagged




File Size [Byte] =
Dimensions
x 3 colours
x Bit Depth x Byte/ 8 Bit

+ size of technical
metadata

List of Image Files in Bridge

Date when picture was taken (not available in Windows)

Give your own keywords

Name	Date Created ↑	Size	Type	Rating	Label	Keywords	Date Modified	Dimensions	Resolution	Color Profile
 TIFF File old.tif	27.07.2015, 12:27	1.42 MB	TIFF image			e-rara File for Workshop	16.11.2020, 18:00	860 x 308	300 ppi	2050021_A_15_10_2014_04
 JPEG File.jpg	22.08.2015, 19:58	855 KB	JPEG file				22.08.2015, 19:58	1600 x 1200	72 ppi	sRGB IEC61966-2.1
 TIFF File without extension	16.11.2020, 17:04	1.42 MB	Binary document				16.11.2020, 18:00			Untagged

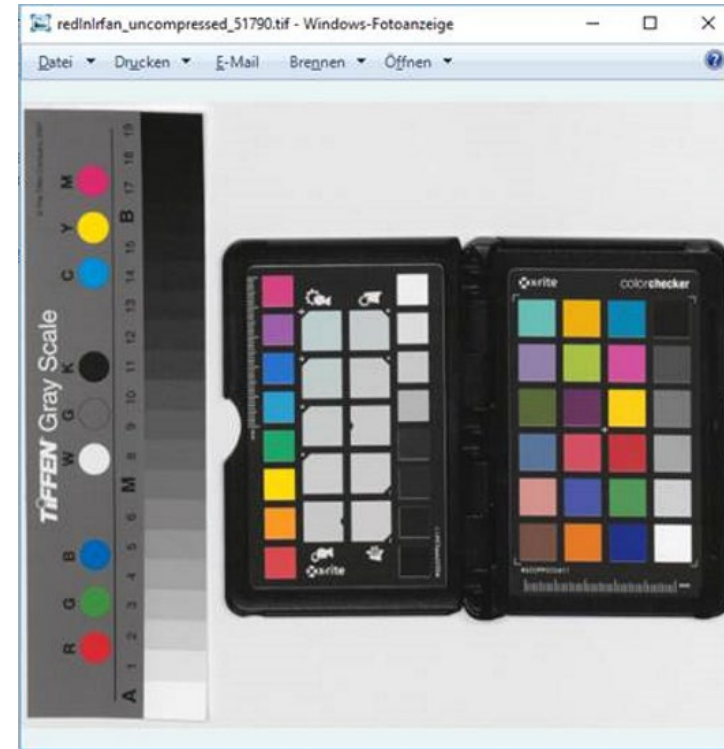
Colour Profile corrects the colours

Colour Profile corrects colours of a scanner or camera (*.tif or *.jpg)

Uncorrected colours



Corrected colours



- Scanner stores the numbers for this correction as ICC Profile or Colour Map (white balance)
- Each image gets a copy in its technical metadata

Problem: Different Viewers show Different Colours (*.tif or *.jpg)

Show uncorrected colours: IrfanView, Windows Media Player, Paint, Paint 3D, XnView



Correct colours: Windows Foto, Foto, Windows Thumbnails, Adobe Acrobat und Adobe Photoshop



See also https://github.com/mpv-player/random-stuff/tree/master/icc_profiles (access date Nov 8, 2021)

Questions ?

Defect files

Bit flips

- Example of simulated bit flip by changing just 1 value in a hex-editor



CRC-Error

- This TIFF files have been compressed as ZIP
- When decompressing there was an error:

1 unterbrochene Aktion

Aufgrund eines unerwarteten Fehlers können Sie die Datei nicht kopieren. Wenn der Fehler weiterhin ausgegeben wird, können Sie mithilfe des Fehlercodes in der Hilfe nach diesem Problem suchen.

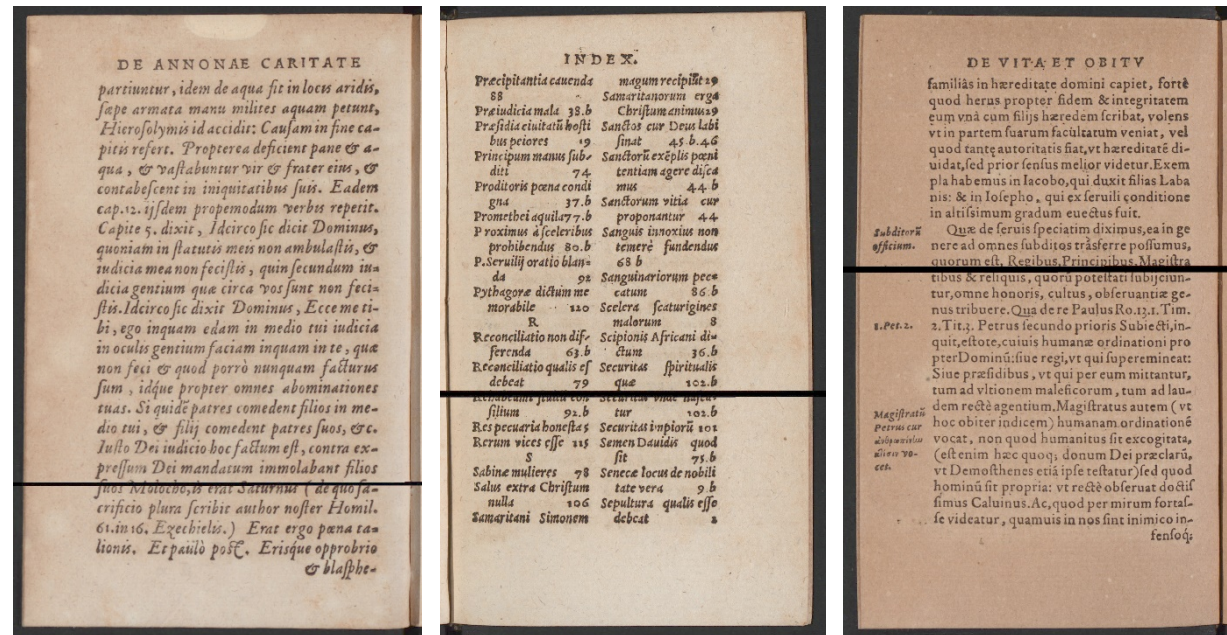
Fehler 0x80004005: Unbekannter Fehler

1320029.tif
 Typ: TIF-Datei
 Änderungsdatum: 22.01.2011 01:59
 Größe: 5.95 MB

Für alle aktuellen Elemente wiederholen

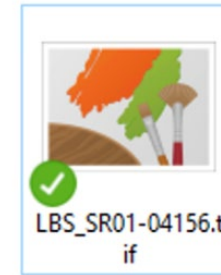
Vorgang wiederholen Überspringen Abbrechen

- Reason: Defect RAM




EOF Error

- Thumbnail is not showing up in Explorer
- It can't be opened neither with Windows nor with Photoshop



Adobe Photoshop

 Konnte "LBS_SR01-04156.tif" weil ein unerwartetes EOF (Ende der Datei) aufgetreten ist nicht öffnen.

OK

- Solution: set correct offset in a hexeditor:

```
Offset (h) 00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F
00000000 49 49 2A 00 74 9F BF 00 13 00 FE 00 04 00 01 00  II*.tÿ¿...p.....
```

```
Offset (h) 00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F
00000000 49 49 2A 00 08 00 00 00 13 00 FE 00 04 00 01 00  II*.....p.....
```

A photograph of a modern library interior. The scene shows a long, bright study area with a long white table and several chairs. A person is seated at the table, working on a laptop. The room has large windows on the right side, providing natural light. The ceiling is high with a skylight. The walls are light-colored, and there are some red accents on the left side. A red banner is overlaid on the image, containing text.

Thank you for your attention

<https://www.umfrageonline.ch/s/a13b937>

Roland Suri
Fabian Schneider

data-archive@library.ethz.ch

ETH Library
Rämistrasse 101
8092 Zurich

<https://library.ethz.ch/en/publishing-and-archiving/archiving/digital-long-term-preservation>