

Private Information in Public Data - A Forensic Analysis of Unaffected User Image Metadata Provided by Online Platforms

Felix Fischer , Elisabeth Ziesch, Dirk Labudde 

Faculty Applied Computer Sciences and Biosciences
Hochschule Mittweida University of Applied Sciences
Mittweida, Germany

e-mail: {fischell | eziesch | labudde}@hs-mittweida.de

Abstract—Metadata provides additional information to an image about its origin or use case. This paper examines how online platforms treat metadata in images. Pictures in the graphic formats of Joint Photographic Experts Group (JPEG), Portable Network Graphics (PNG), Tagged Image File Format (TIFF) and WebP are prepared with numerous metadata using Exchangeable Image File Format (EXIF), Extensible Metadata Platform (XMP) and International Press Telecommunications Council (IPTC). The entire dataset were uploaded and downloaded to different platforms and then evaluated. Forensically relevant metadata, such as copyright, Global Positioning System (GPS) information, camera details and dates, as well as timestamps were examined more closely. This showed that the platforms differed greatly from one another in how they deal with metadata in images. It has not been possible to recognize a uniform handling of metadata on the inspected platforms. The removal of metadata on platforms aligns with the expectations of user privacy.

Keywords-social media; Privacy; Metadata Publishing; Messaging.

I. INTRODUCTION

Every day, millions of users interact with a massive influx of online images through views, ratings and shares. However, visual perception only scratches the surface. Beyond the visual content, pictures could contain hidden private data. Without careful management, publishing these images risks accidentally leaking personal details, such as your identity or exact coordinates. The metadata is shared together with the images without the user being aware of it. With regard to dating platforms, users do not expect to disclose sensitive data. Harassment by stalkers could be an unwanted consequence for users. This can be avoided by thoughtful handling of images by the platforms.

However, this is not the case with photo-sharing platforms, where it is important for users to publish camera-specific information and settings along with the image. Users also do not expect an altering of image data, while using a storage platform. Therefore, a download of a stored image should result in unchanged information. The image should not be distinguishable from the original.

Another perspective arises from the forensic interest in digital image evidence. Information that may be included in the image material provides important data and clues for forensic investigations. Metadata could reveal the location, the photographer, used equipment, used programs, light intensity of the environment and times. This may lead to new subjects or crime scenes.

This study intends to examine the automatic removal or retention of metadata in images on different platforms. For this purpose, metadata is attached to the images, which are then uploaded and downloaded across multiple platforms. The resulting evaluation shows whether online platforms retain, remove or alter metadata in published images and whether users' privacy is compromised. Additionally, the received data can be seen as an upper limit for forensic usage.

The remainder of the paper is organized as follows. Section II explains the formats of image metadata. Section III presents use cases in a forensic context. Section IV describes the methods used in this paper. Section V lists the platforms selected for examination. Section VI presents the findings, and finally, Section VII discusses possible reasons for deviations in our findings.

II. FORMATS OF METADATA IN IMAGES

Metadata is additional attached data to data. It provides secondary information to the stored data to help organizing or understanding the content. They are essential for understanding the data, its storage, preservation and finding data for future use [1].

The Exchangeable Image File Format (EXIF) is a standard format for image metadata developed by Japan Electronic and Information Technology Industries Association (JEITA) [2]. The metadata of an image file contains a variety of information, such as the date and time it was created, all camera settings, and information about the shooting environment, camera manufacturer and model, ISO sensitivity, and horizontal and vertical resolution of the image. Where appropriate, geographical information may also be included. This data is created by the camera or program and then stored in the header of the digital image file. The main purpose of EXIF metadata is to enable precise search, retrieval and viewing of images. This simplifies management and organization, and allows the tracking of the original settings and operations on an image [3]. For example, EXIF metadata can be found in Joint Photographic Experts Group (JPEG), Portable Network Graphics (PNG), Tagged Image File Format (TIFF) and WebP graphic formats [4].

The Extensible Metadata Platform (XMP) format is a standard Adobe format that was introduced in 2001 along with Adobe Acrobat version 5.0 [5]. Based on the Extensible Markup Language (XML) programming language, the metadata format can also be embedded in JPEG, PNG, TIFF, and

WebP image file types [6]. It is format-independent and extensible [2]. Primarily, it is used for image processing and gives many possibilities to describe images extensively and flexibly [7]. Such information can include, for example, copyright information, title and descriptions of the image, as well as geographical information, date and time of creation of the image [1].

The International Press Telecommunications Council (IPTC) is a metadata format, which is gradually being replaced by the metadata format XMP [8]. This format is mainly used in journalistic areas and for museum or cultural heritage collections to store information about the author and the content of the image [2]. This includes contact information, title, description, location, creation date and time. This format can be saved in graphic formats, such as JPEG, PNG and TIFF [9].

III. FORENSIC VALUE OF METADATA IN IMAGES

An image itself can be very useful in a crime investigation. Images containing metadata can provide additional context or information about an author, place or creation process. These additional data could be a date or time. Times can enable an ordering in time of images and therefore confirm or deny alibi. Depending on the sun's position in the sky, the casted shadows change as well during the day. By combining time with shadow directions and length, a geographic location will be possible along a longitude [10].

Metadata about the camera equipment can lead to used hardware to create the image. This indicator could also lead to the author of an image. By analyzing the focus length, a distance to the image objects or measurements inside the image can be taken [11]. The exact location of the used camera during the creation of the image will be possible.

Metadata can also contain information about the shooting environment. Lighting of the surrounding effects shutter-speed and ISO sensitivity. Using this, time or weather could be guessed. In combination with weather reports, a location could be roughly determined.

Global Positioning System (GPS) enables a precise localization worldwide. If the GPS location is included in the metadata, a precise positioning of the camera during the creation of the photo can be deduced. This method was used to locate John McAfee in Guatemala after he fled Belize, where a warrant of arrest for murder was issued [12]–[14].

The copyright field in the metadata informs about the applied rights to the image. They can also contain contact information or names. Therefore, these could be used to determine the author of an image. Additionally, a title or description can be provided. The used language can give hints for the location or nationality of the author.

Sometimes programs override fields in the metadata. This could be used to exclude computers as an origin. If a culprit has not installed the used program, he is likely not the author of the image. Furthermore, can the listing of a program in the metadata be a hint for image manipulation.

IV. METHODS

Writing metadata to images refers to adding additional information, such as author, date of capture, camera details, and GPS information. Different programs can be used to add and read metadata from images [15]. We use the website IMGonline.com.ua to fill the metadata fields with information in a self created image, showing the second building of our university. The included EXIF editor offers many different functions for managing metadata in images. EXIF, XMP, and IPTC metadata in JPEG images can be viewed, edited, copied, or completely removed in tabular form. For EXIF metadata, there are 59 different metadata fields, such as Software, Model, Last Modified Date, Author, ISO, Flash, Color Space, Exposure Mode, Zoom Factor, Focal Length, Contrast, Saturation, Sharpness, Lens Model and Copyright. There are also several fields to edit for IPTC and XMP metadata. Figure 1 illustrates the full process chain. In the first step, we use IMGonline.com.ua to add as much metadata to the image as possible.

However, these functions can only be applied to JPEG images. In order to convert the metadata in the images to other graphic formats, these must be converted. The conversion of the JPEG image into the graphic formats TIFF and WebP can be done without loss of metadata via websites that are freely accessible on the Internet. When converting to the graphic format PNG, however, all metadata is lost, which is why the image editing software GNU Image Manipulation Program (GIMP) was used for conversion [16]. Consequently, the JPEG and TIFF graphic formats contain metadata in EXIF, XMP and IPTC formats, PNG metadata in EXIF, XMP and partially IPTC formats, and the WebP graphic format contains metadata in EXIF and XMP formats.

The dataset is then uploaded to the online platforms listed in Section V. Problems may occur, such as errors, when logging in or verifying the account with the platform, generally uploading the record, and problems with the respective graphic formats or file sizes.

We continue by downloading the images just uploaded. Many of the platforms do not offer the possibility to download published images via a download function. To get around this, the "Save As" function of the browser can be used.

After the dataset has been uploaded and downloaded on different platforms, the images are evaluated. Here, the images of the dataset are compared with the images after the upload and download. The different graphic formats are differentiated. When comparing, objective attention is paid to whether and how the metadata in the images has changed.

The software Manja Digital Asset Management is used for the evaluation [17]. It allows to manage, exchange and archive different file types, such as images, audios and videos. With regard to the metadata, Manja Digital Asset Management offers the possibility to automatically recognize and display important image information. In order to compare the images with each other, two images can be selected with the help of this software. The metadata of the two images displayed in tabular form can

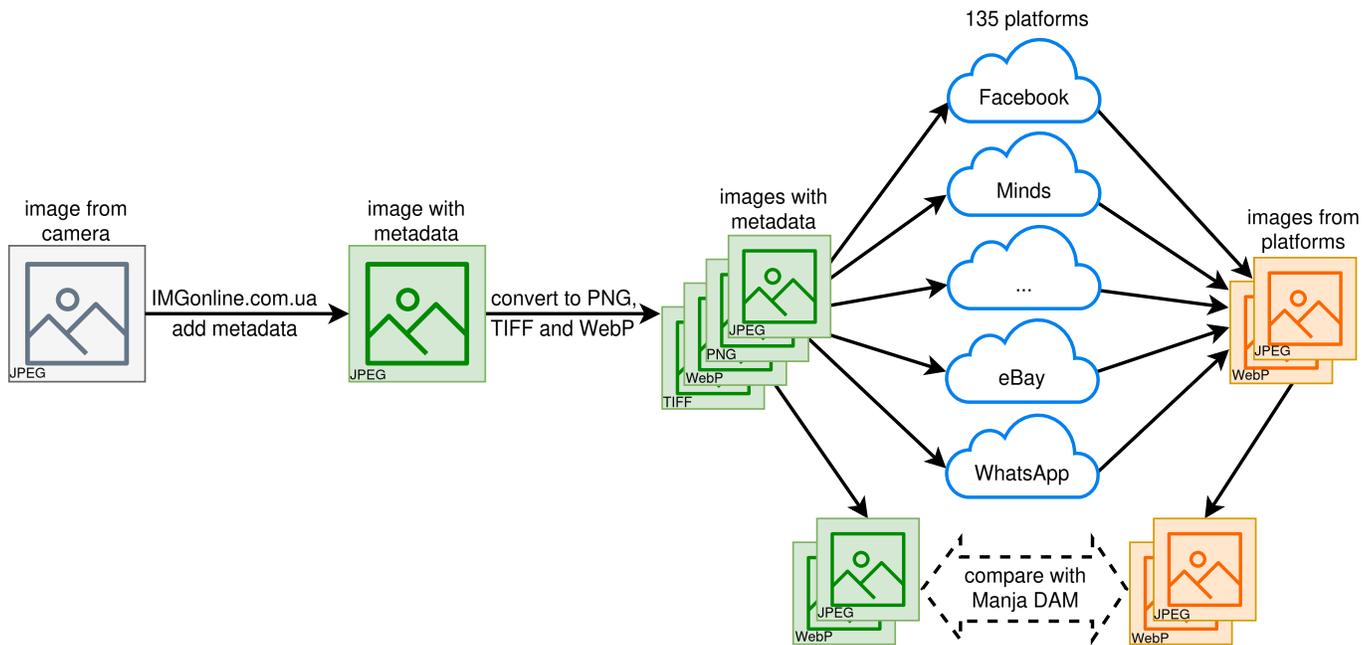


Figure 1. Images are enriched with metadata and compared with the downloaded versions of these images from different platforms. This process is done with the image formats JPEG, PNG, WebP and TIFF. The comparison is realized with the online tool Manja Digital Asset Management.

then be compared - identical information is displayed in black, while non-identical information is highlighted in red.

V. SELECTED PLATFORMS

The selected platforms for the up and download were divided into different categories: social media, community platforms, dating platforms, photo sharing platforms, marketplaces, messengers, cloud storage and cooking platforms. The selection of these categories allows a wide range of online platforms. These platforms do not represent all platforms that exist in this category. It is not possible to take into account all available platforms due to their scope. Below, the online platforms are listed in their respective categories.

Social Media: Amino, ASKfm, Basistar, Befilo, CloutHub, Demoxia, Facebook Lite, Facebook, FC2, Flickr, Gab, Galleria, GORF, hi5, Instagram, LinkedIn, Mastodon, MeWe, Mixi, Pinterest, Plurk, Reddit, Snapchat, SpinChat, Taringa, TikTok, Tumblr, Twitch, Vero, Wer kennt wen, Whisper, X (formerly known as Twitter), Xing, Yareny, Yooco, YouTube

Community Platforms: 23snaps, AboutMe, Academia, Brainly, Care2, CaringBridge, Crockes, DailyStrength, Diaspora, Elpha, Evernote, girls ask guys, GoodReads, Influencer, LiveJournal, Medium, Minds, NextDoor, Opportunity, Patient Like Me, Polywork, Quanswer, Quora, RallyPoint, ResearchGate, Shutterfly, Slack, Snapfish, Stack Overflow, Travellerspoint, Wattpad, Yelp

Dating Platforms: Badoo, BlackPlanet, Bumble, Canoodle, EliteSingles, LOVOO, Match, MeetMe, OkCupid, Tagged, Zoosk

Foto Sharing Platforms: 500xp, ArtStation, Behance, Crevado, DeviantArt, Dribbble, EyeEm, FotoCommunity, GuruShots, Houzz, ImageShack, Imgur, Pexels, PicsArt, Pixaby, PostImage, Raverly, SmugMug, TinyPic, Unsplash, VSCO, YouPic

Marketplaces: Craigslist, eBay, Etsy, Facebook Marketplace, Kleinanzeigen (formerly known as eBay Kleinanzeigen), Shpock, Vinted

Messengers: Discord, Element, Fluffy Chat, Google Chat, HalloApp, ICQ, Line, Messenger (Meta), Messenger von Google, Signal, Skype, Telegram, Viber, VK, Webex by Cisco, WeChat, WhatsApp, Zoom

Cloud Storage: Amazon Photos, Dropbox, Google Drive, Google Photos, Microsoft OneDrive

Cooking Platforms: Chefkoch, Cookpad, Tastemade, Yummly

VI. RESULTS

In total, the dataset was applied to 135 platforms. Of these, 82 removed all metadata from the uploaded images. We list them in Section VI-A. In addition, 19 platforms retained all available metadata associated with the uploaded images. These platforms are listed in Section VI-B. The remaining 34 platforms did not uniformly handle the metadata in the images. Section VI-C explains the handling in more detail. Due to the diversity of the outcomes, we analyze the data from different information perspectives. In Section VI-D we present our findings for the most used platforms in detail. Our findings for platforms that process very sensitive data are presented in Section VI-E. Finally, in Section VI-F we summarize the results.

A. Complete removal of metadata

The analysis shows that many platforms completely remove the metadata in the images after uploading and downloading. Therefore, the images after downloading are free of metadata regardless of the graphic format and form. The image formats used are JPEG, PNG, TIFF and WebP in the form of images, profile pictures and banners. In addition, it was found that both the size of the files and the resolution were reduced.

The following platforms remove the metadata completely: Amino, 23snaps, AboutMe, ArtStation, ASKfm, Badoo, Baisstar, Befilo, BlackPlanet, Brainly, Bumble, Canoodle, Care2, Chefkoch, CloutHub, Cookpad, Craigslist, Crevado, Crockes, DailyStrength, Demoxia, Diaspora, eBay, EliteSingles, Etsy, Facebook Lite, FotoCommunity, Gab, GORF, GuruShots, HalloApp, hi5, Houzz, Imgur, Influenster, Instagram, Kleinanzeigen (formerly known as eBay Kleinanzeigen), Line, LinkedIn, LOVOO, Mastodon, Match, MeetMe, Messenger von Google, Minds, Mixi, NextDoor, OkCupid, Opportunity, Pexels, Pixaby, Plurk, Polywork, Quora, RallyPoint, Reddit, Shpock, Signal, Snapchat, SpinChat, Stack Overflow, Tagged, Tastemade, Telegram, TikTok, Travellerspoint, Twitch, Unsplash, Vero, Vinted, VK, Wattpad, WhatsApp, Whisper, X (formerly known as Twitter), Xing, Yareny, Yelp, YouPic, YouTube, Yummly, Zoosk.

B. Full retention of metadata

The analysis showed that some platforms retain the metadata of the image completely after uploading and downloading. Images in JPEG, PNG, TIFF, and WebP formats were used as images or profile images. Metadata, such as EXIF, XMP and IPTC have been fully preserved. With the exception of two messengers (*Element* and *FluffyChat*), the GPS information remained in the image. The file size and resolution of the images have not changed significantly. With the exception of four cases where the file size or resolution was reduced, the values remained unchanged.

The following platforms retain the metadata completely: FC2, Academia, Amazon Photos, DeviantArt, Dropbox, Element, Evernote, Flickr, Fluffy Chat, Google Chat, Google Drive, Google Photos, Microsoft OneDrive, Quanswer, SmugMug, Snapfish, TinyPic, Webex by Cisco, Zoom.

C. Partial preservation of metadata

The evaluation yielded distinctive results regarding the partial removal of the metadata. Due to their complexity, different scenarios have been created in which important metadata fields are presented and their occurrence on the respective platforms is explained. The main focus for these results is metadata.

Consideration regarding date and time: Information, such as date and time, can be included in images in different forms as metadata. Images can include the date of creation, date of modification, and date of last modification of the metadata including time. This information can be found mainly in the standard metadata as well as in EXIF, XMP and IPTC.

Providing date and time as metadata can be meaningful. In this way, images can be sorted chronologically and the

exact time of capture can be determined. It may also play an important role in criminal cases from a forensic point of view, as it may serve as evidence of certain events or offences. It provides information about whether or not a person was actually present on a certain day and at a certain time. This information may also help in the case of contradictory witness statements, in order to verify the accuracy of the statements. In general, the information serves to support the credibility of digital image evidence.

The following platforms provide date and time information: Behance, CaringBridge, Discord, Elpha, EyeEm, Galleria, girls ask guys, GoodReads, ICQ, ImageShack, LiveJournal, MeWe, Patient Like Me, Pinterest, PostImage, Raverly, ResearchGate, Shutterfly, Skype, Taringa, Viber, Wer kennt wen, Yooco.

Image formats are not treated equally.

Consideration with regard to copyright: The information about the author can be found in the EXIF, XMP and IPTC metadata. There, the name of the author or creator of the image is noted. With the help of this information, the legal distribution and use of the image can be clearly regulated. Images with copyright notice are protected by copyright and allow a clear identification of the author. These images may also help to track the distribution of an image on the Internet.

The following platforms preserve copyright information: Behance, CaringBridge, Discord, Dribbble, Elpha, EyeEm, Facebook, Facebook Marketplace, Galleria, girls ask guys, GoodReads, ICQ, ImageShack, LiveJournal, Medium, Messenger (Meta), MeWe, Patient Like Me, Pinterest, PostImage, Raverly, ResearchGate, Shutterfly, Skype, Taringa, Tumblr, Viber, Wer kennt wen, Yooco.

Image formats are not treated equally.

Consideration in relation to the color space: The color space is crucial information for the printing of images. It determines how the colors are displayed in the image and has a significant influence on the print quality and color accuracy. For example, there are the Red Green Blue (RGB) and Cyan Magenta Yellow Key (CMYK) color spaces. In terms of forensic meaning, the color space provides potential information on whether images have been manipulated or not by comparing them with the original color spaces.

The following platforms keep color space information available: 500px, Behance, CaringBridge, Discord, Elpha, Galleria, girls ask guys, GoodReads, ICQ, ImageShack, LiveJournal, MeWe, Patient Like Me, PicsArt, PostImage, ResearchGate, Shutterfly, Skype, Taringa, Tumblr, Viber, VSCO, Wer kennt wen, Yooco.

Image formats are not treated equally.

Consideration of location information: Information about the location where the image was taken may be included in the metadata of an image. This data makes it possible to determine the exact location, as longitude, latitude and altitude can be specified. The location information can be automatically embedded in the image or added manually. This information is extremely valuable from a forensic point of view, as it can provide information about the location of different images or the reconstruction of movements. The metadata can be used to

confirm that the image was actually taken at the appropriate location and thus determine the exact location of the event. It can also prove or refute inconsistencies in the statements of persons prosecuted with criminal law about their whereabouts. They can also serve as an alibi by being able to confirm the location of a person at a certain time.

The following platforms provide location information: Behance, CaringBridge, Discord, Elpha, Galleria, girls ask guys, ICQ, ImageShack, LiveJournal, MeWe, Patient Like Me, PostImage, ResearchGate, Shutterfly, Skype, Taringa, Wer kennt wen, Yooco.

Image formats are not treated equally.

Consideration in relation to the camera model: The camera model with which the image was taken is stored in the EXIF metadata of the image. There is information about the brand and model of the digital camera, such as "Canon EOS M200" or "Panasonic Lumix DCGX880". *Canon* is the brand and *EOS M200* is the model. The camera model can be of great importance for forensic purposes, in particular to determine the origin and integrity of captured images. The suspect claims to have certain camera equipment, but the metadata of the images shows a different camera model. This may indicate a false statement or manipulation, or may confirm that certain images were actually taken with the specified camera. This can confirm the credibility of the source.

The following platforms retain camera model information: Behance, CaringBridge, Discord, Elpha, Galleria, girls ask guys, GoodReads, ICQ, ImageShack, LiveJournal, MeWe, Patient, PostImage, Raverly, ResearchGate, Shutterfly, Skype, Taringa, Viber, Wer kennt wen, Yooco.

Image formats are not treated equally.

Consideration to the description of the image: The description of an image is included in the EXIF metadata. These descriptions should reflect the content of the image clearly and precisely, that is, explain the main elements and actions depicted in the image. This information can be embedded directly into the metadata of the image or in other cases added as a text field next to the image. The description of the image plays an important role in the accessibility of the Internet. It captures only the image content and can also serve as alternative text for visually impaired people. This ensures that they are able to fully perceive visual content. The image content is clearly communicated, allowing it to be perceived by a larger number of people.

The following platforms preserve information about the description of the image: Behance, CaringBridge, Discord, Elpha, Facebook, Facebook Marketplace, Galleria, girls ask guys, GoodReads, ICQ, ImageShack, LiveJournal, Messenger (Meta), MeWe, Patient Like Me, PostImage, Raverly, ResearchGate, Shutterfly, Skype, Slack, Taringa, Tumblr, Viber, Wer kennt wen, Yooco.

Image formats are not treated equally.

Protection of images from modification: When images are published on the Internet, some information is partially changed, such as file size, resolution, or metadata in the image itself. However, it is often advantageous to leave the images

used unchanged and to maintain their original state. This is important, for example, for photo-sharing platforms where it is important for users to publish camera-specific information and settings along with the image. It is important to keep the metadata in the images.

Protection of the image content itself from modification is given by the following platforms: Academia, Amazon Photos, Behance, CaringBridge, DeviantArt, Discord, Dropbox, Element, Elpha, Evernote, FC2, Flickr, Fluffy Chat, Galleria, girls ask guys, Google Chat, Google Drive, Google Photos, ICQ, LiveJournal, MeWe, Microsoft OneDrive, Patient Like Me, PostImage, Quanswer, ResearchGate, Shutterfly, Skype, SmugMug, Snapfish, TinyPic, Webex by Cisco, Wer kennt wen, Yooco, Zoom.

Image formats are not treated equally.

Protection of personal data: Data protection and personal security require careful handling of personal information in images. To avoid disclosing sensitive location data or copyright information, platforms should remove metadata from published images. While such information is useful for forensic investigations, it does not protect personal rights. With regard to dating platforms, users do not expect to disclose sensitive data.

Platforms do not allow the option of retaining metadata or not. However, the removal may occur automatically if images are converted to other graphic formats without the user being informed.

Protection of personal data is provided by the following platforms: 23snaps, AboutMe, Amino, ArtStation, ASKfm, Badoo, Basistar, Befilo, Behance, BlackPlanet, Brainly, Bumble, Canoodle, Care2, CaringBridge, Chefkoch, CloutHub, Cookpad, Craigslist, Crevado, Crockes, DailyStrength, Demoxia, Diaspora, Dribbble, eBay, EliteSingles, Etsy, Facebook Lite, FoToCommunity, Gab, GORF, GuruShots, HalloApp, hi5, Houzz, Imgur, Influenster, Instagram, Kleinanzeigen (formerly known as eBay Kleinanzeigen), Line, LinkedIn, LiveJournal, LOVOO, Mastodon, Match, MeetMe, Messenger (Meta), Messenger von Google, MeWe, Minds, Mixi, NextDoor, OkCupid, Opportunity, Patien Like Me, Pexels, PicsArt, Pinterest, Pixaby, Plurk, Polywork, PostImage, Quora, RallyPoint, Reddit, ResearchGate, Shpock, Shutterfly, Signal, Slack, Snapchat, SpinChat, Stack Overflow, Tagged, Tastemade, Telegram, TikTok, Travellerspoint, Tumblr, Twitch, Unsplash, Vero, Viber, Vinted, VK, Wattpad, WeChat, WhatsApp, Whisper, X (formerly known as Twitter), Xing, Yareny, Yelp, Yooco, YouPic, YouTube, Yummly, Zoosk.

Image formats are not treated equally.

No conversion of graphic formats: Many platforms convert uploaded images to consistent graphic formats for a variety of reasons. On the one hand, the conversion removes some sensitive metadata from the images in order to protect privacy. On the other hand, platforms require certain requirements for graphics formats. Therefore, the images are converted accordingly to make them compatible. This creates uniformity in the graphics formats used and the image quality can be

adapted to the requirements of the platforms. It also improves the loading times on websites.

The following platforms do not convert images to other graphic formats: 500xp, Academia, Amazon Photos, ASKfm, Basistar, CaringBridge, Crockes, DeviantArt, Diaspora, Discord, Dropbox, Element, Elpha, Evernote, FC2, Flickr, Fluffy Chat, Google Chat, Google Drive, Google Photos, hi5, ICQ, LiveJournal, Mastodon, MeetMe, MeWe, Microsoft OneDrive, Patient Like Me, Pixaby, RallyPoint, Raverly, Skype, Slack, SmugMug, Snapfish, Stack Overflow, TinyPic, Travellerspoint, Unsplash, VSCO, Webex by Cisco, Xing, Yooco, Zoom.

D. Forensic value of popular platforms

Continuing, the ten most popular platforms by active user count in our dataset will be examined in detail [18].

Facebook: It is possible to upload images in the data format JPEG, PNG, TIFF and WebP. Regardless of the used image format, downloaded images will only be available as JPEG. Metadata stored using the XMP standard will be removed completely. However, copyright, author and description will remain in images using the EXIF standard. While utilize the IPTC standard, images will keep information about copyright. The image contains additionally metadata of the e-mail field in IPTC if the uploaded image uses JPEG, PNG or TIFF as image format.

Instagram: Images can be uploaded using JPEG or PNG format. However, PNGs will be converted to JPEGs as well. Images do not contain any metadata after download. Therefore, no forensic value can be expected in the metadata of images provided by Instagram.

WhatsApp: WhatsApp accepts uploads of images as JPEG, PNG, WebP. Regardless of the uploaded format, JPEG will be chosen for download. A conversion happens for other formats to JPEG. The Messenger removes all metadata in the images. Therefore, no forensic value can be expected in the metadata of images of WhatsApp.

YouTube: On YouTube, the profile image was tested. Here is an upload possible as JPEG, PNG and WebP. Downloading the image, it will be converted to JPEG. The resulting image will not contain any metadata at all. Therefore, no forensic value can be expected in the metadata of profile images of YouTube.

TikTok: JPEG, PNG and WebP can be chosen as image format for uploads. But downloaded images will always use JPEG as image format. A conversion happens during the upload or download process. Images provided by the platform will have all metadata removed. Therefore, no forensic value can be expected in the metadata of images from TikTok. Additionally, alters TikTok the image by adding a gray bar at the top and a TikTok watermark.

WeChat: PNG, JPEG, and WebP have been uploaded to WeChat. Downloading results in a conversion to JPEG. Only when using the JPEG image format, the information about the resolution remains in the EXIF metadata. Everything else will be deleted. The resolution is redundant to inspecting the image data itself, so no additional value is provided. Therefore,

no forensic value can be expected in the metadata of images provided by WeChat.

Telegram: Downloaded images on Telegram will always use the JPEG format regardless if JPEG, PNG or WebP are used as the original image format. All metadata will be removed. Therefore, no forensic value can be expected in the metadata of images provided by Telegram.

Messenger (Meta): The messenger from Meta supports uploads with JPEG, PNG, WebP and TIFF format. In downloading these images, a conversion to a different format can happen. TIFF will be converted to JPEG and WebP becomes GIF. Using the EXIF format, copyright and description are included in the metadata. While using IPTC copyright and the E-Mail are available. All fields in the XMP metadata will be deleted. Like in the conversion to a different image format, WebP handles metadata different as well. Images uploaded as WebP will have all metadata removed. Therefore, the available metadata depends on the used image format. Copyright information could lead to the creator of the image.

Snapchat: The upload of images to Snapchat was possible using the image formats JPEG, PNG and TIFF. Images provided by the platform will always use the image format JPEG. These images hold no metadata. Therefore, no forensic value can be expected in the metadata of images from Snapchat.

Reddit: The platform Reddit accepted uploads using JPEG, PNG and WebP as image formats. However, the used image format, images provided by Reddit will always use JPEG. A conversion is done when using PNG or WebP image format. The provided images contain no information in the metadata anymore. Therefore, no forensic value can be expected in the metadata of images provided by Reddit.

E. Platforms with sensitive private data

Analyzing the data, some platforms handle sensitive private data or data intended to share with a small target group. Here, dating platforms and cloud storage providers stand out. The results from the most popular platforms will be presented in detail.

LOVOO: LOVOO allows upload only as JPEG. Consequently, downloads use JPEG as the format as well. No format conversion takes place. LOVOO removes all metadata to protect users from stalkers. Therefore, no forensic value can be expected in the metadata of images from this platform.

Badoo: PNG and JPEG are used as image formats for uploads. However, PNGs will be converted to JPEG. Badoo alters the image content by adding a Badoo watermark. The dating platform Badoo removes all metadata for the protection of its users. Therefore, no forensic value can be expected in the metadata of images provided by Badoo.

Bumble: Upload using the JPEG or PNG image format is possible. Regarding of the used format, a downloaded image will use WebP. Strangely, an upload as WebP is not allowed. Consequently, a conversion is mandatory. The image has no metadata remaining. Users are protected regarding the metadata. Therefore, no forensic value can be expected in the metadata of images from Bumble.

BlackPlanet PNG and JPEG can be used as image formats for uploading to the dating platform BlackPlanet. Downloaded images will always use PNG as the image format. Received images will not contain any metadata. Additionally, the images downloaded using our method are just black. It is the only platform we could observe doing that. Therefore, no forensic value can be expected in images from BlackPlanet.

Amazon Photos: The image cloud storage of Amazon accepts all image formats used in this analysis. The downloaded images will be identical to the uploaded ones. Therefore, all images will contain the metadata available during the upload process. Consequently, images of this platform can yield high forensic value.

Dropbox: This cloud storage provider accepts all image formats used in this analysis. The downloaded images will be identical to the uploaded ones. Therefore, all images will contain the metadata available during the upload process. Consequently, images of this platform can yield high forensic value.

Google Drive: This cloud storage service from Google accepts all image formats used in this analysis. The downloaded images will be identical to the uploaded ones. Therefore, all images will contain the metadata available during the upload process. Consequently, images of this platform can yield high forensic value.

Google Photos: The image cloud storage from Google accepts all image formats used in this analysis. The downloaded images will be identical to the uploaded ones. Therefore, all images will contain the metadata available during the upload process. Consequently, images of this platform can yield high forensic value.

Microsoft OneDrive: This cloud storage service accepts all image formats used in this analysis. The downloaded images will be identical to the uploaded ones. Therefore, all images will contain the metadata available during the upload process. Consequently, images of this platform can yield high forensic value.

F. Summary

The process of uploading and downloading images in different graphic formats with a lot of metadata on different platforms is intended to illustrate how the respective platforms handle metadata. It was found that most platforms remove the metadata from the images, regardless of the graphic format. Out of a total of 135 platforms, this corresponds to a share of 60.74%. Of these, about 25.07% of the platforms retain some of the metadata in the images, depending on the graphic format. The smallest proportion, about 14.07% of the platforms, retained the metadata in the images completely.

No similarities in graphic formats could be identified, regardless of whether the metadata was retained or removed. Also, when converting the images to the respective platforms during uploading and downloading, no similarities were found regarding the processing of metadata.

With regard to the platform categories and their handling with metadata, except three categories, no commonalities could

be identified. In the cooking category, it was found that all metadata in the images is completely removed from all platforms. The same applies to dating platforms. In this regard, it is important that all metadata be removed. Users expect that sensitive private information will not be disclosed to other users. The opposite is cloud storage, with all platforms retaining the metadata completely. Here, users intend to download images in the same state as they were uploaded. Changing the metadata is undesirable.

Basically, the majority of platforms remove metadata from published images. This is a positive step towards protecting personal information. Insofar as the data is needed for digital forensics and law enforcement, the results show that the complete and partial removal of the metadata reduces available information. Accordingly, the information from the visual content of the images is more decisive and usable.

The forensic value is therefore minuscule for most platforms and highly dependent on the platform itself. A broad range of information can be expected on cloud storage platforms. The available metadata in this analysis represents an upper limit of available information. Especially copyright information that has been left untouched often will not be filled in by the daily use of non-photo-enthusiastic users. In this case, the field would also not yield any usable information.

VII. CONCLUSION AND FUTURE WORK

In the following, potential sources of error that could have affected the results of this work or could lead to divergent results in comparative future work are identified.

Due to the increasing number of privacy laws and policies, the information provided by each platform varies. The laws in different countries also differ from one another. Therefore, a consistent presentation of the results over the long term cannot be guaranteed. For example, the handling of metadata in images may be specified in the terms of use or privacy policies of the respective platforms.

Furthermore, not all available public platforms were used for this work. Care was taken to make a selection from different areas. During the upload and download process, some platforms found that it is not possible to upload images. As a result, only 135 of the original selected 227 platforms could be used. Therefore, it is not possible to make general statements for the individual platform categories.

There is a need for optimization with regard to the graphic formats. The TIFF format was not compatible when used on different platforms and therefore often could not be used. Using a more appropriate graphic format would lead to more consistent and better results.

The process of uploading images to different platforms is often complex. On some platforms, images must be manually released or uploaded before being uploaded. Therefore, no statements can be made about those platforms, even if they allow the uploading and downloading of images.

Another point is that the process of uploading and downloading for web based platforms was done only by the same browser (in this case: Microsoft Edge). It is possible that the use of other

browsers could lead to different results. For mobile apps, a virtual device used the apps provided by the Google app store. The well-known messenger *WhatsApp* and the well-known social network *Snapchat* make it possible, among other things, to use the application not only as an app on a smartphone, but also as a web application on a computer. However, these web applications of the platforms were not taken into account in this study, which could also lead to altered results.

The dataset was uploaded and downloaded from Germany, Saxony. Therefore, other countries where the uploading and downloading might lead to different results have not been considered.

The images were downloaded immediately after the image were uploaded. Therefore, no statement can be made about how the metadata behaves when downloading older images. It is unclear whether the platforms used to handle metadata in images differently than they do today. Therefore, the results are only relevant for the current time and can not make statements about the previous handling of metadata by the platforms.

As far as messengers are concerned, the images in this study were shared exclusively in private chats. However, there is the option to extend the upload and download process to group chats. In addition, it remains unclear how the metadata in the images behaves when received in direct or group chats.

Many messengers and social networks offer their users the opportunity to share images both as posts and as stories or statuses. For example, *WhatsApp* and *Instagram* allow sharing images with other users in short sequences for 24 hours. This behavior was not taken into account in the present investigation.

Finally, the content of the image could be subject of a future investigation. TikTok for example adds a watermark to the image. Hidden watermarks or other modifications could be made by other platforms as well.

In conclusion, the majority of platforms remove all metadata from images shared publicly on social networks. No similarities could be identified in terms of graphic formats, different conversion processes or platform categories. The handling of metadata is very individual from platform to platform. The removal of metadata on platforms aligns with the expectations of user privacy. A huge violation could not be found.

ACKNOWLEDGMENT

This paper was funded by the European Union and the Free State of Saxony (Germany).



Kofinanziert von der
Europäischen Union



Diese Maßnahme wird mitfinanziert durch
Steuermittel auf der Grundlage des vom
Sächsischen Landtag beschlossenen Haushaltes.

REFERENCES

- [1] R. Grunzke et al., “The masi repository service — comprehensive, metadata-driven and multi-community research data management”, *Future Generation Computer Systems*, vol. 94, pp. 879–894, 2019, ISSN: 0167-739X. DOI: <https://doi.org/10.1016/j.future.2017.12.023>.
- [2] G. N. Hill and M. A. Whitty, “Embedding metadata in images at time of capture using physical quick response (qr) codes”, *Information Processing & Management*, vol. 58, no. 3, p. 102504, 2021, ISSN: 0306-4573. DOI: <https://doi.org/10.1016/j.ipm.2021.102504>.
- [3] H.-C. Huang and W.-C. Fang, “Metadata-based image watermarking for copyright protection”, *Simulation Modelling Practice and Theory*, vol. 18, no. 4, pp. 436–445, 2010, Modeling and Simulation Techniques for Future Generation Communication Networks, ISSN: 1569-190X. DOI: <https://doi.org/10.1016/j.simpat.2009.09.004>.
- [4] P. Harvey, *Exif-tags*, 2026. Accessed: Feb. 2, 2026. [Online]. Available: <https://exiftool.org/TagNames/EXIF.html>.
- [5] N. L. Romero et al., *Recovery of descriptive information in images from digital libraries by means of EXIF metadata*. Emerald Group Publishing Limited, 2008, p. 3. DOI: <https://doi.org/10.1108/07378830810880388>.
- [6] P. Harvey, *Xmp-tags*, 2025. Accessed: Feb. 2, 2026. [Online]. Available: <https://exiftool.org/TagNames/XMP.html>.
- [7] PDFlib GmbH, *Überblick über xmp (overview of xmp)*, 2023. Accessed: Feb. 2, 2026. [Online]. Available: <https://www.pdfli.com/de/pdf-know-how/xmp/ueberblick-ueber-xmp/>.
- [8] T. Hoffmann-Walbeck et al., *Standards in der Medienproduktion (Standards in media production)*. Springer Vieweg Berlin, Heidelberg, 2013, pp. 82–85, ISBN: 978-3-642-15042-5. DOI: <https://doi.org/10.1007/978-3-642-15042-5>.
- [9] P. Harvey, *Iptc-tags*, 2020. Accessed: Feb. 2, 2026. [Online]. Available: <https://exiftool.org/TagNames/IPTC.html>.
- [10] P. Kakar and N. Sudha, “Authenticating image metadata elements using geolocation information and sun direction estimation”, in *2012 IEEE International Conference on Multimedia and Expo*, 2012, pp. 236–241. DOI: 10.1109/ICME.2012.82.
- [11] M. L. Heuschkel, F. Fischer, and D. Labudde, “Objektive bildqualitätsmessung in der digitalforensik : Mtf-basierte quantifizierung der bildqualität und erzielbarer messgenauigkeit für forensische bildanalysen (lenses image quality measurement in digital forensics : Mtf-based quantification of image quality and achievable measurement accuracy for forensic image analysis)”, de, *NextGen Scientific Review*, no. 4, pp. 17–24, 2026, ISSN: 2940-0929. DOI: 10.48446/opus-16388.
- [12] A. Liou, *John mcafee is trapped in paradise*, 2012. Accessed: Feb. 18, 2026. [Online]. Available: <https://www.vice.com/en/article/aee3gj/john-mcafee-silicon-valleys-greatest-villain-is-trapped-in-paradise>.
- [13] Lyle, *Vice.com john mcafee exclusive reveals his location in iphone exif data*, 2017. Accessed: Feb. 18, 2026. [Online]. Available: <https://www.mobileprivacy.org/2012/12/vice-com-publishes-exclusive-with-john-mcafee-reveals-location-in-iphone-metadata-exif/>.
- [14] A. Sawall, *John mcafee mit iphone-geolocation geortet (john mcafee located by iphone geolocation)*, 2012. Accessed: Feb. 18, 2026. [Online]. Available: <https://www.golem.de/news/vice-john-mcafee-mit-iphone-geolocation-geortet-1212-96131.html>.
- [15] F. Fischer, E. Ziesch, and D. Labudde, “Erfassungssoftware für metadaten in bildern (extraction software for metadata in images)”, de, *Polizeiinformatik*, no. 2024, pp. 209–218, 2024.
- [16] GIMP, *Gnu image manipulation program*, <https://www.gimp.org>, 2026.

- [17] R. Frunzke, *Manja digital asset management*, <https://manjadigital.de/>, 2026.
- [18] D. Thuy, *Most popular social networks worldwide as of february 2025, by number of monthly active users*, 2025. Accessed: Feb. 2, 2026. [Online]. Available: <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>.