

Metadaten und Bilder

Christoph Rinne*

09. Januar 2022

Inhaltsverzeichnis

1	Allgemeine Information	1
1.1	Vorbemerkung	1
1.2	Bildformate	2
2	Metadaten	2
2.1	IPTC	2
2.2	EXIF	3
3	Software	3
3.1	XnView/XnViewMP	3
3.2	exiv2 Anwendungsbeispiele	4
3.3	ExiTool Anwendungsbeispiele	5

1 Allgemeine Information

Im Kern geht es um Metadaten und deren Bedeutung für die Nachnutzung von digitaler Objekte. Fotos, oder allgemeiner Rasterbilder, sind hier nur das gewählte Medium, um mit allgemein bekannten Objekten auf dieses wichtige Thema nachdrücklich hinzuweisen.

1.1 Vorbemerkung

Bilder - und es geht hier um Rasterbilder bzw. Fotos - haben durch die digitale Fotografie einen enormen Umfang erreicht, diese stellen in der Anzahl und vor allem im Datenvolumen bei fast allen (Forschungs-) Projekten den größten Bestand dar. Zwar gibt es eine immer bessere Bildanalyse und mit trainierten Neuronalen Netzen werden heute erstaunliche Ergebnisse erzielt (z.B. [YOLO](#)). Dennoch macht es Sinn, die bekannten Informationen zu den Bildern nicht (nur) in getrennten Listen zu führen, sondern in die Bilder hineinzuschreiben. Mit Blick auf die Nachnutzung zu dem vorgenannten noch das Problem der unzähligen Bildformate: von den RAW-Dateien der jeweiligen Kamera über JPG, TIFF und PNG als den bekanntesten zu weiteren Software spezifischen Formaten und Varianten innerhalb der Formate.

Metadaten finden Sie in vielen anderen Objekten, auch in Word und PDF. Bei MS Word finden Sie einige Standardfelder mit Datei > Vorbereiten > Eigenschaften und zudem unter Dokumenteneigenschaften > Erweitert die Möglichkeit, weitere Felder einzubinden. Nach dem Speichern als PDF finden Sie diese in der Regel unter Datei > Eigenschaften. Beim Drucken in ein PDF werden diese Metadaten nach meiner Kenntnis nicht übernommen.

Automatisch eingebundene Metadaten in allen digitalen Dokumenten können viele Informationen zum Urheber der Datei transportieren. Sie sollten diese Daten also bewusst einsetzen und kontrollieren.

Leider kann über Metadaten auch Schadcode transferiert werden ([stackoverflow](#)).

*Christian-Albrechts Universität zu Kiel, crinne@ufg.uni-kiel.de

1.2 Bildformate

Rasterbilder sind in ihrer Grundstruktur einfach, sie entsprechen einer Matrix in der jeder Zelle ein Wert zugewiesen wird, z.B. von 0 bis 255 (Schwarz bis Weiß). Ein Bild kann mehrere Ebenen besitzen (RGB) oder auch eine Farbtiefe von mehr als 256 (28) Farbwerten besitzen (224, true color). Derart gestaltete Bilder werden schnell sehr groß, weshalb diverse Kompressionsalgorithmen entwickelt wurden (JPG oder TIFF mit LZW). Im Fall von JPG werden dabei die Bilder unwiederbringlich verändert und qualitativ schlechter. Eine andere Strategie besteht in der Indexierung der wirklich verwendeten Farben (Farbpaletten-Modus, PNG). Eine gute Einführung und Überblick in das Thema Formate und Archivierung sowie weiterführende Links bietet [IANUS-FDZ](#). Als geeignet für die Archivierung gelten unkomprimierte TIF, GeoTIF, DNG. Wegen der sehr eingeschränkten Metadaten ([W3Org](#)) gilt PNG nur bedingt als geeignet und JPG gilt wegen der Kompression als ungeeignet. Wegen der hohen Verbreitung und der deutlich geringeren Größe ist JPG aber *de facto* ein Standard geworden. Um die Qualität von JPG nicht unnötig zu verschlechtern sollten sie wiederholtes Speichern vermeiden und bevorzugt mit verlustfreien Funktionen, z.B. der Rotation um 90° arbeiten. Auch das Schreiben von Metadaten in JPG erfolgt verlustfrei.

2 Metadaten

Bei Fotos gibt es zwei wichtige und schon lange existierende Gruppen von Metadaten mit jeweils definierten Feldern: IPTC und EXIF. IPTC zielt eher auf Urheberrechte und Bildinhalt, während EXIF mehr technische Informationen umfasst. Einige Felder, u.a. Datumsangaben, sind dabei redundant vorhanden. Bei XMP handelt es sich um eine allgemeine auf XML Standard basiertes Schema zur Definition von Metadaten. XMP wird allgemein verwendet, für PDF und Bilderformate.

Weitere Informationen zu dem Thema Metadaten finden Sie im Bereich Projektphasen > Dokumentation der IT-Empfehlungen von [IANUS](#)

2.1 IPTC

IPTC steht für *International Press Telecommunications Council* und den Datenstandard, der zusammen mit der *Newspaper Association of America (NAA)* als *Information Interchange Model (IIM)* entwickelt wurde. Einen guten Überblick und Einstieg bietet die [Erläuterung zu dem Programm exivtool](#). Trotz der Definition kommen unsauber geschriebene Daten vor (u.a. Länge der Felder, Speicherort des IPTC-Blocks). IPTC Daten werden in Blöcken (*envelope records*) organisiert und sind innerhalb dieser erneut nummeriert (*tag id*). Beachten Sie bitte, es gibt ein IPTC Core Schema und ein IPTC Extension Schema. Eine kleine Auswahl der oft genutzten Felder aus den *IPTC Application Record Tags* sehen sie hier:

tag id	tag name	writable
80	By-line	string[0,32]+
85	By-lineTitle	string[0,32]+
90	City	string[0,32]
92	Sub-location	string[0,32]
95	Province-State	string[0,32]
100	Country-PrimaryLocationCode	string[3]
101	Country-PrimaryLocationName	string[0,64]
103	OriginalTransmissionReference	string[0,32]
105	Headline	string[0,256]
110	Credit	string[0,32]
115	Source	string[0,32]
116	CopyrightNotice	string[0,128]
118	Contact	string[0,128]+

Nicht alle Tags werden von jeder Software verwendet oder deren Bezeichnung sogar angemessen oder identisch übersetzt. Eine kleine Auswahl aus älteren Programmversionen sehen Sie hier:

IrfanView 3.97	exiv2 0.18.2	photoshop elements 3.0	XnView 1.96.2
Copyright	Copyright	CopyrightInformationen	CopyrightVermerk
Caption	Caption	Beschreibung	Objektbeschreibung
Caption writer	Writer	Verfasser der Beschreibung	Verfasser
Headline	Headline		Überschrift
Special instructions	SpecialInstructions		Sonstige Hinweise
Keywords	Keywords	Stichwörter	Sichwörter
Category	Category		Kategorie
Supplemental categories	SuppCategory		Zusätzliche Kategorien
Urgency	Urgency		Dringlichkeit
Byline	Byline	Autor	Fotograf/Ersteller
Byline title	BylineTitle		Titel des Fotografen/Ersteller
Credits	Credit		Bildrechte
Source	Source		Quelle
Object name	ObjectName	Dokumenttitel	Objektname
Date created	DateCreated		Erstellt am (JJJJMMTT)
City	City		Stadt
ProvinceState	ProvinceState		Bundesland
Country	CountryName		Land/Staat

2.2 EXIF

Die Bezeichnung steht für *Exchangeable Image File Format*, es handelt sich überwiegend um technische Metadaten zur Aufnahme, unter anderem Kameramodell, Zeitpunkt der Aufnahme und Kameraeinstellungen. Hervorzuheben sind hier die **Geokoordinaten** und weitere Angaben zu den Verwaltungseinheiten (von Land bis Ort/Stadt). EXIF-Daten werden u.a. in TIF und JPG (JFIF) geschrieben. Bei letzteren kann es durch konkurrierende Angaben zur Position innerhalb der Datei zu Konflikten beider Standards kommen. Die von Mobiltelefonen gerne abgelegten Positionsdaten können in Kombination mit den Hardwareinformationen zum Gerät sehr viel zum Urheber bekannt geben. Sie sollten sich dessen stets bewusst sein.

3 Software

Es sollen drei Programme zum Lesen und Schreiben von Metadaten in Rasterbildern hier genannt werden:

- **XnView /XnView XMP**: Beide Programme stehen als freeware für Win, Mac und Linux zur Verfügung. Sie bieten eine umfassende Bildverwaltung mit umfangreichen Optionen zum Lesen und Schreiben von Metadaten.
- **exiv2**: Ist eine freie, plattformunabhängige und konstant weiterentwickelte Software zum Lesen und Schreiben von Metadaten auf der Kommandozeile. Sie kommt u.a. in der Grafiksoftware Gimp und darktable zum Einsatz. Die Software ist in der aktuellen Version unter Windows auf die vorherige Installation von MinGW oder Cygwin angewiesen, um eine Linuxumgebung zu simulieren und die benötigten Bibliotheken bereitzustellen (*.dll).
- **ExifTool**: Eine Alternative zu exiv2 die ebenfalls auf der Befehlszeile genutzt wird. Da es mir um die Darstellung eines automatisierten Prozesses auf der Kommandozeile geht, dazu ausführlich exiv2, wird nachfolgend nur sehr minimalistisch vorgestellt.

3.1 XnView/XnViewMP

Für dieses Programm (XnView) existiert ein umfangreiches englisches [Wiki](#). Dort finden Sie auch eine Anleitung und ein Handbuch [XnView \(Windows\) User Guide](#). Es ist weniger ein Programm zur Arbeit mit Metadaten von Bildern als vielmehr ein umfassendes Programm zur Verwaltung von Bildern. Das impliziert aber auch das Schreiben und das Lesen von Metadaten (Werkzeuge > Metadaten). Auch hier können *Reguläre Ausdrücke* verwendet werden.

3.2 exiv2 Anwendungsbeispiele

Das ausführen der Befehle auf der Kommandozeile bietet die Möglichkeit, viele Fotos mit spezifisch konfektionierten Befehlen auf einmal abzuarbeiten. Diese Befehle können aus einer Datenbank oder mit einem Tabellen-Programm (z.B. MS Excel) generiert werden. Das nachfolgende Beispiel soll nur einen Einstieg erleichtern, auf der Homepage steht ein umfangreiches [Handbuch](#) zur Verfügung, aus dem die Nachfolgenden Beispiele entnommen wurden. Die Befehlsstruktur sieht wie folgt aus:

```
exiv2 [ opt [arg] ]+ [ act ] file ...
```

Nach dem Befehl folgt die Option, worunter hier u.a. -e (-extract), -i (-insert) und -M (-modify) verstanden werden. Hierauf folgt das auszuführende Argument aus Feld und Feldwert. Abschließend wird die zu ändernde Datei benannt. Eine Verkettung in einem Befehlsaufruf ist möglich:

```
exiv2 -M"add Iptc.Application2.Credit String Mr. Smith" -M"add Xmp.myprefix.Whom  
Mr. Smith" -M"set Exif.Image.Artist Mr. Smith" image.jpg
```

Auf der Homepage von exiv2 finden Sie:

- weitere [Beispiele](#),
- eine [Liste der Metadaten](#) und
- die Liste der in [exiv2 verwendeten IPTC Datensätze](#).

Für das nachfolgende Beispiel laden Sie sich bitte den hier bereitgestellten Ordner /foto/ mit den beiden Bildern und das Programm exiv2 (s.o.) herunter. Speichern Sie für diesen Einstieg alles in einem Ordner "foto" auf Ihrem Desktop um lange Pfadangaben zu vermeiden. Achten Sie auch später bitte auf kurze Ordnernamen ohne Leerzeichen.

Danach rufen Sie die Kommandozeile (Terminal) auf und wechseln in den Bildordner. Unter Windows 10: <Win>+<s>, nach "cmd" suchen und "Eingabeaufforderung" ausführen. Sie sollten mit der Kommandozeile in Ihrem Nutzerverzeichnis gelandet sein: C:\Users\<name>. Hier wechseln Sie auf den Desktop und in den Bildordner: `cd Desktop\foto`.

Als Hilfe einige grundlegende Befehle der Kommandozeile von Windows:

- dir: directory / list, Inhalt des Ordners
- cd: change directory, Ordner wechseln
- . von der aktuellen Position aus
- .. aus dem übergeordneten Ordner
- *.jpg alle Dateien mit der Endung .jpg
- help: eine Liste wichtiger Befehle
- dir /? Hilfe zu dem Befehl dir

Führen Sie nachfolgend in der Kommandozeile bitte folgende Befehle aus:

```
exiv2.exe pr *.jpg  
exiv2.exe -M "add Iptc.Application2.Writer String Name des  
Verfassers von 'Byline'" *.jpg  
exiv2.exe -M "add Iptc.Application2.Credit String Name des  
Fotografen" *.jpg  
exiv2.exe -M "add Iptc.Application2.Byline String Fotos aus dem  
Kieler Umland." *.jpg
```

Sie müssen diese Befehle ggf. natürlich modifizieren, 1. durch ergänzende Pfadangaben zu exiv2 und den Bildern als auch 2. bei dem jeweiligen Inhalt, als 'Name...' und 'Fotos aus...'.

Ein Beispiel für die Stapelverarbeitung mit Excel finden Sie in der Datei: iptc-write.xlsx. Diese Datei hat vier Spalten:

1. vollständige Dateiname (0001-0001-001.jpg)
2. Bildautor
3. Bildinhalt
4. Spalte mit der Anweisung die Daten der Spalten A-C zu einem Befehl zu verkett(=verkett()):

=VERKETTEN("exiv2.exe -M" "add Iptc.Application2.Credit";B2; " " -M" "add Iptc.Application2.Writer";
B2; " " -M" "add Iptc.Application2.Caption"; C2; " " " ;A2)

Anmerkung: Da das " eine besondere Bedeutung in Excel für das Auszeichnen von Zeichenketten hat muss es hier mit einem " auskommentiert werden. " " " fügt also den Text "<Leerzeichen> ein.

Daraus wird dann:

```
exiv2.exe -M"add Iptc.Application2.Credit Christoph Rinne" -M"add Iptc.Application2.Writer Christoph Rinne" -M"add Iptc.Application2.Caption Neolithisches Megalithgrab bei Birkenmoor" 0001-0001-001.jpg
```

Die 4. Spalte enthält abschließend alle Befehle für die Änderungen abhängig von den Zellen in den Spalten A-C. Diese Spalte kann als ganzes in die Zwischenablage kopiert und auf der Befehlszeile eingefügt werden. Dies ist eine Möglichkeit, lange Listen mit Bildinformationen in die Fotos zu schreiben. Machen Sie vorher Backups von den Fotos und prüfen Sie nachträglich die Lesbarkeit (s.o.).

3.3 ExifTool Anwendungsbeispiele

Im Unterschied zu exiv2 handelt es sich bei [ExifTool](#) für Windows um eine einzige ausführbare Datei. Die heruntergeladene Datei sollte von 'exiftool(-k).exe' in 'exiftool.exe' umbenannt werden. Das Programm kann auf der Befehlszeile genutzt werden: `exiftool -a -u -g1 -w <bildname>` oder `exiftool -a -u -g1 -w <bildname> > info.txt` (Geduld!). Daneben können Bilddateien oder ganze Ordner im Explorer von Windows auch durch "ablegen" auf dem Programm bearbeitet werden. Die notwendigen Parameter müssen dann aber durch eine Umbenennung des Programmes vorab übergeben werden, z.B. "exiftool(-a -u -g1 -w).exe". Weiter Informationen durch aufrufen von exiftool.exe in der Kommandozeile ohne Parameter oder auf der [Homepage](#).