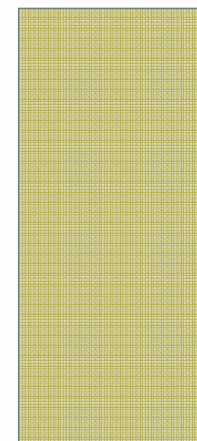


IPTC UND KEYWORDING

BVPA SEMINAR 30. NOVEMBER 2012



TEIL I. IPTC STANDARDS

METADATEN IM BILD

WAS IST IPTC EIGENTLICH?

- Die IPTC (International Press Telecommunication Council) ist eine internationale Organisation, die sich mit der Standardisierung von Nachrichtenformaten beschäftigt.
- Die Standards der IPTC berücksichtigen immer sowohl technische als auch inhaltliche Aspekte.
- Eine Reihe dieser Standards betreffen auch die inhaltliche Beschreibung von Bildern.

TECHNISCHE GRUNDLAGEN

- Die auf die Bildbeschreibung zutreffenden IPTC Standards IIM und IPTC Core beinhalten u.a. die Beschreibung in welcher Form und ggf. an welcher Stelle, die Textinformationen zum Bild (Metadaten) in einer Bilddatei eingebettet sein sollen.

TECHNISCHE GRUNDLAGEN

- Die Daten nach dem alten IIM Standard folgen einem Datenfeld Prinzip. Für jeden Inhalt wird ein virtuelles Datenfeld in bestimmter Größe und an einer bestimmten Stelle in der Bilddatei vorgehalten. Nicht ausgefüllte Felder bleiben vorhanden, sind aber leer.
- Die Beschreibung nach IPTC Core setzen technisch auf Adobe XMP auf und folgt im Großen und Ganzen einem XML Ansatz. Die Daten können in beliebiger Reihenfolge stehen und werden nur durch einen festgelegten Bezeichner definiert. Der gesamte Textblock wird durch Anfangs- und Endmarkierung umschlossen.

TECHNISCHE GRUNDLAGEN

- IIM und XMP sind aber nicht die einzigen Metadaten, die einem Bild mitgegeben werden können. Die beiden wichtigsten weiteren sind:
- EXIF (Exchangeable Image File Format)
- PLUS (Picture Licensing Universal System)

TECHNISCHE GRUNDLAGEN / EXIF

- EXIF wurde von der Organisation der japanischen Kamera Hersteller (Jeita) entwickelt und bietet die Möglichkeit, bereits im Moment der Aufnahme Metadaten automatisch im Bild zu hinterlegen.
- Dieser Standard ist recht flexibel und so werden von Hersteller zu Hersteller abweichende Informationen gespeichert. In der Regel sind dieses z.B.: Zeit und Datum, Objektiv, Brennweite und ähnliche Informationen aber tlw. auch schon GPS Koordinaten.

TECHNISCHE GRUNDLAGEN / PLUS

- Der PLUS Standard wurde von der gleichnamigen Organisation aus Bildanbieter und Bildnutzer Verbänden entwickelt.
- Inhaltlicher Schwerpunkt ist die Übermittlung von erworbenen Nutzungsrechten (PLUS Codes) und Urheber-Informationen. Hierzu gibt es auch eine von PLUS betriebene Registrierungsstelle in der weltweit eindeutige Urheberkennungen vergeben werden.

INHALTLICHE GRUNDLAGEN / IIM

- Der IIM (Information Interchange Model) Standard ist der älteste (letzte Überarbeitung Version 4.1 von 1999), aber derzeit noch am weitesten verbreitete Standard zur Bildbeschreibung.
- Er definiert gut 20 Datenfelder zur Beschreibung der Bildurheber und Bildinhalte.
- Die wichtigsten sind:

ÜBERSICHT ÜBER DIE IIM FELDER

IPTC Feldname	Beschreibung
Object Name	Interner Name des Bildes. Text oder Nummer oder beides
Urgency	Dringlichkeit, ein Wert von 1 (niedrig) - 8
Original Transmission Reference	Die Bildnummer die zum Kunden hin kommuniziert wird
By-Line (Author)	Fotograf
By-Line(Author position)	Stellung des Fotografen (frei, angestellt o.ä.)
Credit	Name desjenigen der das Geld bekommt
Source	Quelle des Bildes, kann identisch mit Credit sein
Copyright Notice	Vollständiger, gewünschter Fotonachweis
Writer/Editor	Autor der Bildbeschreibung
Contact	Kontaktadresse der Bildquelle (z.B. die eigene Agentur)
Special Instructions	Hinweise zur Bildnutzung an den Kunden (keine Werbung, nur Deutschland etc.)

ÜBERSICHT ÜBER DIE IIM FELDER

IPTC Feldname	Beschreibung
City	Ort der Aufnahme
Province/State	Bundesland / Bundesstaat der Aufnahme
Country/ Primary Location Name	Land der Aufnahme (ausgeschrieben)
Country/ Primary Location Code	Land der Aufnahme ISO Code (z.B. GER)
Sublocation	Örtlichkeit innerhalb der Stadt der Aufnahme oder auch ein geografischer Ort ohne Stadtnamen (Rathausplatz oder Grand Canyon)
Date Created	Datum der Aufnahme (Ortszeit)
Headline	Kurzbeschreibung des Bildinhaltes
Caption	Beschreibung des Bildinhaltes
Keywords	Ergänzende Stichworte zum Bildinhalt
Category	Kategorisierung des Bildes unter redaktionellen Gesichtspunkten. Z.B. POL für politisches Thema. Erleichtert automatische Zuordnung der Bilder nach Redaktionen
Supplemental Categories	Weitere Kategorisierung, verfeinert oder ergänz das Vorhergehende. Z.B. WI (Wirtschaft) Handel, Einzelhandel

INHALTLICHE GRUNDLAGEN / IPTC CORE

- Der IPTC Core ist die zeitgemäße Überarbeitung des IIM
- Eine Vielzahl neuer Datenfelder macht ihn noch flexibler als seinen Vorgänger, aber auch umfangreicher (ca. 80 Datenfelder)
- Verbesserte Bezeichnung des Urhebers
- Bessere Details für die Weitergabe strukturierter Daten
- Teilweise Einbindung von Exif und PLUS

ÜBERSICHT ÜBER NEUE IPTC CORE FELDER

IPTC Core Feldname	Beschreibung
Location Shown in the Image	Diese Felder betreffen nur Bilder, auf denen ein geografischer Ort abgebildet ist.
<ul style="list-style-type: none"> • World Region 	
<ul style="list-style-type: none"> • Country Name 	
<ul style="list-style-type: none"> • Country ISO-Code 	
<ul style="list-style-type: none"> • Province or State 	
<ul style="list-style-type: none"> • City 	
<ul style="list-style-type: none"> • Sublocation 	
Location created	Diese Feldgruppe ist für den Ort der Aufnahme gedacht, die Unterteilung ist wie oben.
Event	Bezeichnung eines Ereignisses, z.B. Olympiade
Name of Organisation Featured in the Image	Vollständiger Name einer Organisation die im Bild abgebildet ist (Sozialdemokratische Partei Deutschlands)
Code of Organisation Featured in the Image	Entsprechend hier die Abkürzung (SPD)

ÜBERSICHT ÜBER NEUE IPTC CORE FELDER

IPTC Core Feldname	Beschreibung
Person Shown in the Image	Name der abgebildeten Person (journalistische Fotos)
Model Age	Alter des Modells zum Zeitpunkt der Aufnahme
Minor Model Age Disclosure	Eventuelle, durch das Alter begründete Nutzungsausschlüsse (z.B. Tabak oder Alkohol bei Minderjährigen)
Model Release Status	Existiert ein Model Release?
Model Release Id	Ggf. die agenturinterne Nummer des Releases
Property Release Status	Besteht ein Property Release?
Property Release Id	Ggf. die interne Nummer davon
Artwork or Object in the Image	Eine Gruppe von Feldern zum Beschreiben von abgebildeten Kunstwerken oder anderen Objekten
<ul style="list-style-type: none"> • Title 	Titel des Kunstwerkes
<ul style="list-style-type: none"> • Creator 	Schöpfer des Kunstwerkes
<ul style="list-style-type: none"> • Date Created 	Herstellungsdatum des Kunstwerkes
<ul style="list-style-type: none"> • Copyright Notice 	Ggf. zusätzlicher Urheberhinweise
<ul style="list-style-type: none"> • Source 	Quelle (z.B. ein Museum / eine Sammlung)
<ul style="list-style-type: none"> • Source Inventory Number 	Inventarnummer der Quelle (nicht die der Agentur)

IIM / IPTC CORE KOMPATIBILITÄT

- Diese Unterschiedlichkeit in Technik und Inhalt führt leider dazu, dass die beiden Standards nicht miteinander kompatibel sind.
- Ein Bild kann Metadaten in jedem der beiden Standards enthalten und auch beide nebeneinander.
- Bildbeschriftungs-Software muss dieses dadurch ausgleichen, dass sie gleichzeitig beide IPTC Versionen liest und schreibt.
- Dieses wird derzeit nur zum Teil geleistet.
- Inhaltlich ist IIM der kleinste gemeinsame Nenner.

TEIL II. KEYWORDING

DER WEG DURCH DEN BILDERDSCHUNDEL

KEYWORDING

- Unter Keywording (Verschlagwortung) versteht man im allgemeinen die den Freitext ergänzende Beschreibung durch Stichwörter / Schlagworte (engl. Keywords).
- Für Verschlagwortung gibt es eine Vielzahl von Standards. Dieses führt im Ergebnis dazu, dass es keinen Standard gibt. Zumindest keinen, der sich in der Bildbranche durchgesetzt hat.

KEYWORDING

- Im Gegensatz zu wissenschaftlichen Stichwortlisten und Thesauri, sind Bildinhalte nicht nur sachlich sondern häufig auch assoziativ beschreibbar.
- Die von der IPTC entwickelten Kategorien helfen nur bedingt, da sie im Wesentlichen auf journalistische Bilder zugeschnitten sind.
- Eine weitere Herausforderung bei der Verschlagwortung ist, diese in einer Form abzufassen, die auch ungeübte Rechercheure schnell zum richtigen Bild gelangen lässt.

KEYWORDING

- Zielrichtung einer klassischen Verschlagwortung ist die Auffindbarkeit von Bildern in Datenbanken, sowohl durch Fachleute als auch durch Nichtfachleute (Kunden).
- Es soll ermöglicht werden, Bilder so präzise wie nur eben möglich aufzufinden, wobei die Treffermenge möglichst alle gesuchten Bilder enthält ohne zu viele Fehltreffer anzuzeigen.

KEYWORDING

- Am sichersten erreicht man dieses durch
 - Zusammenfassung von Gruppen gleichthematischer Bilder durch Verwendung von Gruppenbezeichnungen
 - Beschreibung des Inhaltes des jeweiligen Einzelbildes mit den zutreffenden Suchbegriffen
 - Abstraktion des einzelnen Bildinhaltes im Bezug auf z.B. eine symbolische Verwendung
 - Verwendung eines kontrollierten Vokabulars

KEYWORDING

- Je heterogener das Bildangebot und damit auch die Zielgruppe, desto problematischer die Bildbeschreibung.
- Was für den einen Kunden lediglich eine Rose ist, ist für den Redakteur eines Fachbuches eine *Rosa tomentosa*.
- Ebenso sollte eine Spezialagentur, die im Wesentlichen wissenschaftliche Bezeichnungen verwendet, überlegen, ob mit zusätzlichen umgangssprachlichen Begriffen nicht der Kundenkreis erweitert werden kann.

KEYWORDING / KEYWORD SPAMMING

- Oft drängt sich bei dem einen oder anderen der Gedanke auf, viele Keywords = viele Treffer.
- Dieses ist nur insofern zutreffend, dass die verwendeten Keywords auch zum Bild passen müssen.
- Wird hierbei zu weit über das Ziel hinausgeschossen, kann es dazu führen, dass der Kunde in einem Übermaße Bilder präsentiert bekommt, die absolut nichts mit seiner Suche zu tun haben und er sich dann sehr schnell wieder verabschiedet.

KEYWORDING / KEYWORD SPAMMING

- Als abschreckendes Beispiel sei an dieser Stelle ein Bild eines schlichten Kochtopfes auf neutralem Untergrund mit einem schräg sitzenden Deckel genannt, das ca. 100 Stichworte enthält u.A. Trennungsjahr, ungenießbar, Pärchen, Menschen, Liebe, Mediation, Partnervermittlung, Problem, Leidenschaft, Gemeinschaft, e-Herd, Gerichtsverfahren und noch vieles mehr

KEYWORDING / KONTROLLIERTES VOKABULAR

- Damit innerhalb eines Bildbestandes themengleiche Bilder auch die gleichen Keywords verwenden, benötigt an ein Kontrolliertes Vokabular.
- Unter einem Kontrollierten Vokabular versteht man eine Zusammenstellung festgelegter Begriffe für definierte Sachverhalte.
- Deren Nutzung ist für alle Betextenden verbindlich.

KEYWORDING / KONTROLLIERTES VOKABULAR

- Bei der Entwicklung eines kontrollierten Vokabulars, das man für den eigenen Bildbestand nutzen möchte, sollte man sich im Vorwege einige grundsätzliche Gedanken machen:
 - Aus welchen Themenbereichen biete ich Bilder an?
 - Wer sind meine Kunden, welche Sprache (nicht Nationalsprache) sprechen diese?

KEYWORDING / EMOTIONEN

- Das Beschreiben von Emotionen birgt eine ganze Reihe von Schwierigkeiten.
- Neben wenigen emotionalen Grundstimmungen (Freude, Trauer, Angst, Schmerz, Wut) gibt es eine Vielzahl von Zwischenstimmungen und eine ebensolche Vielzahl von Begriffen.
- Diese sind häufig sehr subjektiv und schwer, eindeutig in Bildern auszudrücken.
- Dementsprechend uneindeutig wird die Zuordnung.

KEYWORDING / EMOTIONEN

- Der Versuch, jede nur mögliche Emotion eines Bildes zu beschreiben, ist dem Kunden gegenüber sicherlich gut gemeint, führt aber m. E. nicht zum Ziel. (Ist ein geschlachteter Fisch nun ekelig oder lecker?)
- Der Kunde erhält den (falschen) Eindruck nach allen möglichen Emotionen suchen zu können und wird so mittelfristig fehlkonditioniert.
- Besser: Der Kunde versucht erst sich das Bild vorzustellen, das er zum Ausdruck der Emotion benötigt und sucht dann nach diesem Bild ohne emotionale Begriffe zu verwenden. Dieses dem Kunden zu kommunizieren, ist Aufgabe des Verkäufers.

KEYWORDING / SPRACHLICHE FORMEN

- Gern werden für Rubrikenbezeichner Pluralformen gewählt, hiervon rate ich aber ab.
 - Ein Kunde der auf der Suche nach dem Bild von einem Baum ist, verwendet in der Regel 1. nicht den Suchbegriff „Bäume“ und möchte 2. üblicherweise kein Bild mit vielen Bäumen finden wenn er nach „Baum“ sucht.
 - Ebenfalls möchte jemand, der nach mehreren Bäumen mit dem Suchbegriff „Bäume“ sucht, nicht seitenweise Bilder mit einem einzelnen Baum finden.
 - Eine korrekte Verwendung von Plural und Singular Formen erhöht die Trefferquote.

KEYWORDING / SPRACHLICHE FORMEN

- Bei Verben sollte, zusätzlich zu der im Freitext verwendeten Form, immer auch der Infinitiv verwendet werden.
 - Steht im Freitext nur „Paar geht spazieren“, dann würde ein Kunde unter den Suchbegriffen „Paar, spazieren gehen“ nicht fündig werden.

KEYWORDING / SPRACHLICHE FORMEN

- Bei jeder Diskussion über Keywording kam früher oder später die Frage nach den „Teekesselchen“ (Homonyme) auf. Also gleiche Begriffe mit unterschiedlicher Bedeutung wie z.B. Bank.
- Meiner Erfahrung nach, wird dieses Problem überbewertet.
 - Der vollständige deutsche Wortschatz umfasst ca. 500.000 Wörter
 - Der umgangssprachliche Wortschatz immer noch ca. 50.000
 - Homonyme gibt es einige hundert

KEYWORDING / SPRACHLICHE FORMEN

- Durch den Kontext beschreibende oder durch gruppierende zusätzliche Stichwörter kann man den Inhalt eingrenzen.
 - Bank + Geld oder Bankhaus, Bankfiliale etc.
 - Bank + Sitzbank und / oder Möbel

KEYWORDING / DATEN

- Ein Kalenderdatum ist in der Regel ein Datum im Format TT.MM.JJJJ
- Es gibt aber auch andere Formen:
 - 2012
 - 21 Jahrhundert
 - 70er Jahre
 - 1970er Jahre
- Je nach verwendeter Datenbank sollte überlegt werden auch diese Formen zu berücksichtigen.

KEYWORDING / FAZIT

- Für die Verschlagwortung von Bildern gibt es keine allgemeingültige Schablone
- Keywords sollen Bilder in Datenbanken schnell und präzise auffindbar machen
- Synonyme und ähnliche Begriffe sollten in die Keywords
- Keywords sollen Begrifflichkeiten verwenden die sowohl intern als auch extern zum Ziel führen