# plainpicture

Uta Kaack, Image Data Manager

# Efficient metadata - use and occuring questions at plainpicture

IPTC Photo Metadata Conference 2015, Warshaw

# plainpicture

- plainpicture was founded in 2001 as an alternative source for commercial imagery with a strong focus on unusual, unexpected and new photography, showing image solutions away from the mainstream
- Hamburg-based agency with offices in London, Paris and New York. plainpicture offers about 750.000 highly edited images 1/3 own photographers, 2/3 partner collections
- More than 1000 international photographers under contract, providing exclusive material
- Collections of more than 90 international partner agencies
- Clients with high expectations of creative imagery advertising, editorial clients, book publishers

### I want to talk about

- Which metadata do we use?
- Metadata in our daily business
- Clients expectations and needs
- Problems with metadata quality and how to deal with them
- Insight into our handling of keywording metadata

# Metadata we use

- Object name / ID
- By-line / Creator
- Date created
- Special instructions
- Place names: City/State or Province/Country
- Headline
- Keywords
- Credit/Copyright information

#### **Release information**

#### **Model release**

- We demand model release information as soon as there are people in the image
- Manual check of every delivered batch of images
- Callback to partner agency or photographer to ascertain the possibility of getting missing releases
- If not successful: exclusion of images
- Original model release is stored by the photographer or the partner agency who assumes responsibility for it. We store the releases of our own photographers
- Sensitive usage: In case of a sensitive usage request we try to contact the model via the partner agency or photographer to check a possible usage

#### **Property release**

The legal situation is extremely different in every country. Partner agencies are responsible for delivering controlled images.

# Metadata in our daily business

- Metadata may be delivered either in IPTC, .csv, .txt or .xls files. We convert all
  metadata into .xls. The control and correction of file names and keywording is done
  in Excel due to its flexibility every partner and every new batch brings particular
  requirements.
- In our internal processing, we don't use IPTC which would be too time-consuming.
   In our systems, we hold every image file more than 10 times (LowRes, HighRes, backup copies, different servers, etc.)
- For all interfaces IPTC is absolutely necessary: photographers upload their image metadata in IPTC, partner agencies deliver or demand IPTC metadata, clients download images with IPTC metadata via Internet.
- Partner keywording data is mapped with plainpicture vocabulary with various possibilities to adjust it
- Only searchable field: keywording
- Title and place names are not search-sensitive because they are not controlled/ controllable

#### Clients needs and business effects

#### Clients needs

- Getting all relevant images for a search without search ballast. Base: concise, relevant and complete keywording
- Geographical information: important for specific searches. Often missing or incomplete.

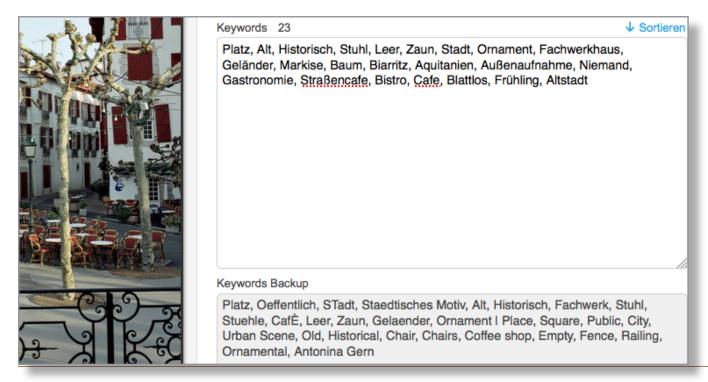
#### **Business**

- Poor data quality means more human input on all levels of the workflow and therefore more costs. But....
- plainpicture policy: as a medium-size agency we cannot enforce our standard to others. We want smaller agencies with interesting images as partners – and accept their sometimes imperfect data quality. We offer support and individual feedback, communicate a lot and sometimes even open our indexing interface for them.

# **Problems with metadata quality**

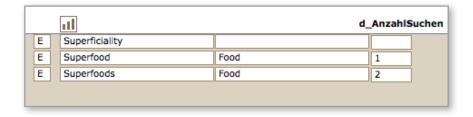
### Bilingual or multilingual keywording

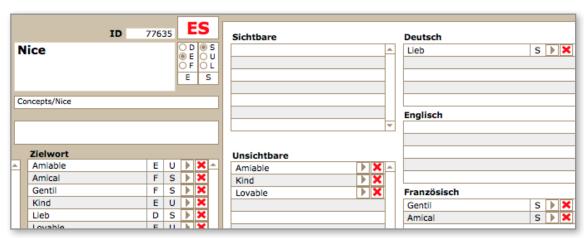
- IPTC neither offers fields for an alternative language in keywording and headline nor a flexible, undefined field which could be used
- Indexing data occurs in more than one language
- Vertical bar is used as a divider between two languages process data means more effort, time and errors



# No separation of languages, two or more languages are mixed up

Nederland, Amsterdam, et	en, diner, smikkelen, smullen, artisjok, tafel, aan tafel, vers, groente, superfood super food, uber groente, gezond kind kindje,
Amb aanhangfiets	fiets, fietsen, bike, bicycle, fietshelm, vader, papa, kinderen, children, ijburglaan, holland, the netherlands
Amb vakantiehuis in italie	vakantie, vacation, voetballen, soccer, lente, groen, polmone
Amb kinderen in italie	umbrie, umbria, italie, italy, kinderen, children, tuinieren, vakantie, vacation, lente, spring, plezier, fun
Amb timetable	vertrektijdentabel, vertrektijden, vertrektijd, opzoeken, kijken, reizen, travel, ov, openbaar vervoer, grootouders, kleindochter
Amb meisje in keuken	kitchen, girl, klutsen, beslagkom, garde, koken, cook, kokkerellen, keukenschort, schort, kind, child, peuter, kleuter, beslag
Amb nul energie woning	huishouden, household, koken, cook, computer, laptop, strijken, ironing, iron, huiskamer, living room, keuken, kitchen, stroon
Amb iPad	apparaat, generatie, generation, apple, tablet, device, kind, child
Amb leren fietsen	ijburg, brug, bridge, fietsen, leren fietsen, bike, child, kind, meisje, girl, oefenen, practice, bikes, bicycle, zwaar trappen, doortr
Amb draaimolen	fair, kermis, fietsen, racefiets, jurk, dress, meisje, girl, merry go round, bike, draaimolen





# Non-native speaking keywording

- Non-valid or unusual English words and expressions used as keywording terms
- Spelling mistakes

Ancient, Antique, Being Made, By Hand, Children's toys, Cottage, Country, Country House, Detail, Door, Door Handle, Door Knob, Doorhandle, Doorknob, Europe, Hung, Idea, Indoors, Inside, Interior, Interior Decoration, Interior Design, Knit, Knitted, Living Area, Make, Making, Manual Work, No One, No-One, Nobody, Nostalgia, Room Door, Rural, Rustic, Scandinavia, Series, Style, Suspended, Tiles, Toy, Traditional, White, Whiteness, Whitish

#### **Effects**

- Loss of information, terms are not translated
- Expanding vocabulary with redundant data
- Time-consuming maintenance of vocabulary and crevision of incoming indexing data



# **Other problems**

- Special characters which are cracked or mixed up
- Keywording placed in several fields or columns

Stillife / Object;Detail;Box;Padlock;Decay;Worn;Locked;Dirty	Spring	Sunny, Weather	Brown, Yellow		Metal
Architecture; House; Building	Spring	Sunny, Blue sky, Weather	Red		Wooden
Architecture House; Building	Spring	Sunny, Blue sky, Weather	Red		Wooden
Architecture; Detail; Maritime; Traditional; Fishermans shed; Fishermans sh	a Spring	Summer, Weather	Grey		Wooden
Stillife / Object;Pier;Jetty;Detail;Maritime;Decay;Worn;Oar;Worn out	Spring	Sunny, Weather	Brown	Coast	Wooden
Stillife / Object; Pier; Jetty; Detail; Maritime; Decay; Worn; Oar; Traditional; W	/ Spring	Sunny, Weather	Brown	Coast	Wooden
Stillife / Object;Pier;Jetty;Detail;Maritime;Decay;Worn;Oar;Traditional;W	/ Spring	Sunny, Weather	Brown	Coast	Wooden
Stillife / Object;Detail;Nail;Decay;Worn	Spring	Sunny, Weather	Brown		Metal
Architecture; Detail; Decay; Worn; Traditional; Log house	Spring	Summer, Weather	Grey		Wooden
Architecture; Detail; Structure / Background; Plank; Wall; Flaking paint; Dec	Spring	Summer, Weather	Brown		Wooden
Architecture;Detail;Form;Flaking paint;Decay;Worn;Rust;Rusty;Mounting	Summer, Weather	Brown		Metal	
Architecture; Detail; Decay; Worn; Log house; Wall	Spring	Summer, Weather			Wooden
Architecture; Pier; Jetty; Maritime; House; Building; Traditional	Spring	Sunny, Blue sky, Weather	Yellow	Coast	Wooden
Stillife / Object; Pier; Jetty; Detail; Maritime; Form; Traditional; Boat; Block; Ro	Spring	Sunny, Weather		Coast	Wooden

## All problems mixed together:

sandal, woman, feet, black and white, in

love, masculine, legs, embracing, hug, humanity, femaleen plein

air,amoureux,brouiller,botte,couple,embrassement,sentiment,féminin,féminité,folklorique,haut,horizontal,jambe,amour,masculin,homme,prðs

de,partie,personnes,pittoresque,rapport,relations,sandale,sentiment,escar pin-sandale,cerf,courroie de sandale,femme,pieds,noir et

blanc,amoureux,masculin,jambes,étroit,étroit,humain,fémininim Freien,amourös,verschwommen,Stiefel,Paar,umarmen,Gefühl,feminin, Weiblichkeit,Volk,hoch,horizontal,Bein,Liebe,mÃ×nnlich,Mann,in der NÃ×he,Teil,Menschen,urig,Beziehung,Beziehung,Sandale,Gefühl,Slingback,Hirsch,Riemchensandale,Frau,Füße,schwarz-

weiß,verliebt,maskulin,Beine,drücken,kuscheln,Menschlichkeit,Weibchenna

zewnä...trz,zakochany,zamazane,but,para,obejmowac,uczucie,kobiecy,kobiecosc,ludzie,wysoki,poziom,noga,miå,osc,mä™ski,mä™å¼czyzna,blisko,czä™sc,ludzie,uroczy,relacja,zwiä...zek,sandaå,,uczucie,sandaå,,samiec,sandaå,y,kobieta,nogi,czarno-

biaÅ,y,zakochany,mÄ™ski,nogi,przytulac,przytulac,czÅ,owieczeÅ,,stwo,sam icaal

fresco,amoroso,difuminar,bota,pareja,abrazando,sentimiento,femenino,fe minidad,pueblo,alta,horizontal,pierna,amor,varón,hombre,cercano,parte,

# Handling keywording metadata

- Dynamic multilingual vocabulary
  - Partner keywording is mapped on plainpicture vocabulary
  - New term entries are checked and integrated in the semantic web, if relevant
- Regular measuring of search logs
  - Criteria for relevance of new terms
  - Observation of upcoming issues and concepts
  - Constant evaluation of preferred terms based on searches
- Migration of irrelevant database entries into a separate list (Blacklist)
  - Deletion of (first) names, numbers, codes and irregular strings
  - Replacement of spelling mistakes
  - Individual maintenance of divergent keywording terms by certain partners

# **Blacklist**

		Transfer of the second of the		1			1	11
E	Nationality		67		511413	14.11.2013	15:41:47	p996
E	Open air	Outdoors	191		511424	14.11.2013	15:45:56	p996
E	Disappear		101		520696	15.11.2013	11:26:18	p996
E	Viewpoint		169		501692	15.11.2013	14:48:38	p996
E	Geology	Geography	145		522867	05.12.2013	10:37:31	p996
E	Geography	Geology	200		447017	05.12.2013	10:39:40	p996
E	Food Industry		49		489976	06.12.2013	16:03:43	p996
E	Nave		103		506881	07.01.2014	12:36:36	p996
E	Rear	Backside	549		503912	15.01.2014	11:40:33	p996
E	Intellectual		249			21.02.2014	15:15:26	p996
E	Elder	Elderly person	302		108544	19.08.2014	12:55:46	p996
E	Walker	Walking	781		84488	05.09.2014	10:56:50	p996
E	Interiors	Indoors	567		110941	15.09.2014	13:43:31	p996
E	Delicacy		121		75958	12.12.2014	15:36:29	p996
E	Insulation		289			21.02.2014	12:50:07	p945
E	Insecure	Eerie	1011			21.02.2014	14:37:18	p945
E	Photography		3413		524510	10.12.2013	16:29:36	p934
E	View point		18			29.12.2011		p924
E	Crop		832			01.03.2012		p924
E	Sayings		2			01.03.2012		p924
E	Anatomy		523			01.03.2012		p924
E	Above		1655			01.03.2012		p924
E	Alone	One	5449			02.03.2012		p924
E	White	Caucasian	9885		358113	02.03.2012		p924

# Online keywording tool for photographers

