

Exiv2 - Bug #1081

Read XMP values from CR2 raw file when stored in XMLPacket

15 May 2015 10:11 - Eric Mesa

Status:	Closed	Start date:	15 May 2015
Priority:	Normal	Due date:	
Assignee:	Alan Pater	% Done:	100%
Category:	xmp	Estimated time:	10.00 hours
Target version:	0.26		
Description			
CR2 file is given tag, title, and description in Digikam: https://www.dropbox.com/s/jvqiinpsaapzxf4/_MG_3694.CR2?dl=0			
Processing is done and a JPEG is produced: https://www.dropbox.com/s/ip503k3tom9ciew/_MG_3694.jpg?dl=0			
When I check with EXIFTOOL, the tag, title, and description are in both the CR2 and the JPEG, however, Digikam (which uses Exiv2 and told me to come here with this bug report), only sees the description, not the title and tag.			
So I installed Exiv2 on my computer and here's the output:			
exiv2 _MG_3694.CR2			
File name : _MG_3694.CR2			
File size : 9180576 Bytes			
MIME type : image/x-canon-cr2			
Image size : 3888 x 2592			
Camera make : Canon			
Camera model : Canon EOS DIGITAL REBEL XTi			
Image timestamp : 2015:05:13 17:08:52			
Image number :			
Exposure time : 1/500 s			
Aperture : F2			
Exposure bias : 0 EV			
Flash : No, compulsory			
Flash bias : 0 EV			
Focal length : 50.0 mm			
Subject distance: 0			
ISO speed : 100			
Exposure mode : Aperture priority			
Metering mode : Partial			
Macro mode : Off			
Image quality : RAW			
Exif Resolution : 1936 x 1288			
White balance : Auto			
Thumbnail : image/jpeg, 8709 Bytes			
Copyright :			
Exif comment : Leaves and Scarlett			
Not sure why it doesn't show the tag and title, but I can see them in Digikam			
Here's the JPEG:			
xiv2 _MG_3694.jpg			
Warning: Ignoring IPTC information encoded in the Exif data.			
Warning: Ignoring XMP information encoded in the Exif data.			
File name : _MG_3694.jpg			
File size : 1962688 Bytes			
MIME type : image/jpeg			
Image size : 3898 x 2594			
Camera make : Canon			
Camera model : Canon EOS DIGITAL REBEL XTi			
Image timestamp : 2015:05:13 17:08:52			
Image number :			
Exposure time : 1/500 s			

Aperture : F2
Exposure bias : 0 EV
Flash : No, compulsory
Flash bias : 0 EV
Focal length : 50.0 mm
Subject distance: 0
ISO speed : 100
Exposure mode : Aperture priority
Metering mode : Partial
Macro mode : Off
Image quality : RAW
Exif Resolution : 3898 x 2594
White balance : Auto
Thumbnail : None
Copyright :
Exif comment : Leaves and Scarlett

Why does it say it's ignoring the IPTC info? Because that's where, I'm pretty sure, the info is stored.

Thank you.

Related issues:

Related to Exiv2 - Bug #900: TIFF images lose XMP packet on write if exiv2 wa...	Closed	13 May 2013
Related to Exiv2 - Feature #992: Better raw file support and test	Assigned	18 Sep 2014
Related to Exiv2 - Feature #922: Add options -pS and -dl to application exiv2	Closed	25 Sep 2013
Related to Exiv2 - Feature #1154: Read XMP and IPTC data from Exif Data in JPEGs	New	15 Jan 2016

Associated revisions

Revision 4184 - 14 Jan 2016 10:14 - Robin Mills

#1081 Added Cr2Image::printStructure()

History

#1 - 15 May 2015 11:29 - Alan Pater

What did you set the tag and title to?

#2 - 15 May 2015 12:26 - Alan Pater

Eric, have you set digikam to "If possible, write Metadata to RAW files (experimental)"?

#3 - 15 May 2015 12:56 - Eric Mesa

Alan Pater wrote:

What did you set the tag and title to?

The tag is "Scarlett"

The title is "Scarlett Playing with Leaves"

Eric, have you set digikam to "If possible, write Metadata to RAW files (experimental)"?

Yes. That's how the data got there to begin with. The only thing that's carrying over is the Exif Comment. Nothing else is carrying over from CF2 to JPEG

#4 - 15 May 2015 13:59 - Alan Pater

It's all in both the cr2 and the jpg, but the XMP is embedded inside Exif ApplicationNotes as shown by exiftool.

```
| 16) ApplicationNotes (SubDirectory) -->
| + [XMP directory, 4868 bytes]
| | XMPToolkit = XMP Core 4.4.0-Exiv2
| | Software = digiKam-4.9.0
| | DateTime = 2015-05-13T17:08:52
| | CreatorTool = digiKam-4.9.0
| | CreateDate = 2015-05-13T17:08:52
| | MetadataDate = 2015-05-13T17:08:52
| | ModifyDate = 2015-05-13T17:08:52
```

```
| | DateTimeOriginal = 2015-05-13T17:08:52
| | DateCreated = 2015-05-13T17:08:52
| | Urgency = 0
| | PickLabel = 0
| | ColorLabel = 0
| | ImageDescription = Leaves and Scarlett
| | UserComment = Leaves and Scarlett
| | TagsList = Scarlett
| | CaptionsAuthorNames =
| | CaptionsDateTimeStamps = 2015-05-13T20:33:02
| | RegionList =
| | RegionInfoRegions =
| | LastKeywordXMP = Scarlett
| | HierarchicalSubject = Scarlett
| | Title = Scarlett Playing with Leaves
| | Description = Leaves and Scarlett
| | Subject = Scarlett
```

#5 - 15 May 2015 14:02 - Alan Pater

digikam Bug 347737 - Tagged RAW files do not show tags after round trip to RawTherapee

https://bugs.kde.org/show_bug.cgi?id=347737

#6 - 15 May 2015 14:21 - Eric Mesa

Alan Pater wrote:

digikam Bug 347737 - Tagged RAW files do not show tags after round trip to RawTherapee

https://bugs.kde.org/show_bug.cgi?id=347737

Alan,

Thanks for tracing things across both bug trackers.

#7 - 15 May 2015 16:24 - Eric Mesa

Based on https://bugs.kde.org/show_bug.cgi?id=347737#c10 in which you commented:

"Ok, it looks like what exiftool calls ApplicationNotes, exiv2 calls Exif.Image.XMLPacket."

What are your thoughts at this point? Where does the issue appear to lie?

#8 - 15 May 2015 17:03 - Alan Pater

- Subject changed from Raw file given tag and title in Digikam. After processing in RawTherapee, resulting JPEG's tag and title not read by Digikam. to Unable to read XMP from CR2 raw file when stored in XMLPacket

- Status changed from New to Assigned

- Target version changed from 0.25 to 0.26

It appears that exiv2 has no method to read XMP metadata when it is stored in Exif.Image.XMLPacket.

#9 - 15 May 2015 17:18 - Eric Mesa

Alan Pater wrote:

It appears that exiv2 has no method to read XMP metadata when it is stored in Exif.Image.XMLPacket.

Bummer that it's for 0.26 as that comes out next year at best.

At any rate, just for completion's sake on the changed subject of the bug - it's actually able to read the CR2 file in Digikam as that's where I tagged it. It's the JPEG that's created where it can't read the data stored in Exif.Image.XMLPacket.

In other words, I tag the files in Digikam. And Digikam can read the tags on the DNGs and CR2s. It's only the created JPEGs that can't be read. Which seems to be weird, based on where it's stored. Or maybe Digikam is doing some extra stuff on the side like a Database or something? It's tough when we have 3+ programs interacting.

Also, this holds for DNGs as well as CR2 raw files.

See:

DNG:

Day-2-Scarlett-running-after-RawTherapee-and-back-in-Digikam-463x300.png

CR2:

Day-2-Scarlett-and-Tree-after-RawTherapee-in-Digikam-lost-title-and-tags-488x300.png

Finally, it's weird that Exiv2 puts the data there, but then can't access it again...or maybe I've misunderstood something subtle.

#10 - 15 May 2015 22:10 - Eric Mesa

For comparison purposes, here are DNG and JPEG files exhibiting the same problem. (Showing that it's not unique to CR2)

DNG:

<https://www.dropbox.com/s/pmz8oneual5pq82/Pots.dng?dl=0>

JPEG:

<https://www.dropbox.com/s/l8rw3zpjgf599f7/Pots.jpg?dl=0>

#11 - 15 May 2015 23:08 - Alan Pater

Let's see:

```
~$ exiv2 -g XMLPacket Pots.jpg
Warning: Ignoring IPTC information encoded in the Exif data.
Warning: Ignoring XMP information encoded in the Exif data.
Exif.Image.XMLPacket          Byte      7246  (Binary value suppressed)

~$ exiv2 -g XMLPacket Pots.dng
Error: Directory Canon with 25665 entries considered invalid; not read.
Exif.Image.XMLPacket          Byte      7246  (Binary value suppressed)

~$ exiv2 -px Pots.jpg
Warning: Ignoring IPTC information encoded in the Exif data.
Warning: Ignoring XMP information encoded in the Exif data.

~$ exiv2 -px Pots.dng
Error: Directory Canon with 25665 entries considered invalid; not read.
Xmp.dc.format                 XmpText    9  image/dng
Xmp.dc.creator                 XmpSeq     1  Eric Mesa
Xmp.dc.rights                  LangAlt    1  lang="x-default" 2014 Eric Mesa; This work is lice
nsed to the public under the Creative Commons BY-NC-SA
Xmp.dc.title                   LangAlt    1  lang="x-default" Pots Title
Xmp.dc.description             LangAlt    1  lang="x-default" Pots Caption
Xmp.dc.subject                 XmpBag     1  pots
Xmp.aux.SerialNumber           XmpText    9  820420518
Xmp.aux.LensInfo               XmpText   17  50/1 50/1 0/0 0/0
Xmp.aux.Lens                   XmpText   12  EF50mm f/1.8
Xmp.aux.LensID                 XmpText    2  29
Xmp.aux.ImageNumber           XmpText    2  48
Xmp.aux.FlashCompensation      XmpText    3  0/1
Xmp.aux.OwnerName              XmpText    9  Eric Mesa
Xmp.aux.Firmware               XmpText    5  1.0.4
Xmp.xmp.ModifyDate             XmpText   19  2015-05-13T17:12:43
Xmp.xmp.CreateDate             XmpText   19  2015-05-13T17:12:43
Xmp.xmp.CreatorTool            XmpText   41  Adobe Photoshop Lightroom 5.7.1 (Windows)
Xmp.xmp.MetadataDate           XmpText   19  2015-05-13T17:12:43
Xmp.xmp.Rating                 XmpText    1  5
Xmp.photoshop.DateCreated      XmpText   19  2015-05-13T17:12:43
Xmp.photoshop.CaptionWriter    XmpText    9  Eric Mesa
Xmp.photoshop.Urgency          XmpText    1  0
Xmp.xmpMM.DocumentID           XmpText   44  xmp.did:f4db08a3-598a-a045-85d8-2cf419e1b5ec
Xmp.xmpMM.OriginalDocumentID   XmpText   32  51711369AC381F3A72B5B60DDF7F26A1
Xmp.xmpMM.InstanceID           XmpText   44  xmp.iid:f4db08a3-598a-a045-85d8-2cf419e1b5ec
Xmp.xmpMM.History              XmpText    0  type="Seq"
Xmp.xmpMM.History[1]           XmpText    0  type="Struct"
Xmp.xmpMM.History[1]/stEvt:action XmpText    7  derived
Xmp.xmpMM.History[1]/stEvt:parameters XmpText   68  converted from image/x-canon-cr2 to image/dng, sav
ed to new location
Xmp.xmpMM.History[2]           XmpText    0  type="Struct"
Xmp.xmpMM.History[2]/stEvt:action XmpText    5  saved
Xmp.xmpMM.History[2]/stEvt:instanceID XmpText   44  xmp.iid:f4db08a3-598a-a045-85d8-2cf419e1b5ec
Xmp.xmpMM.History[2]/stEvt:when XmpText   25  2015-05-13T20:05:29-04:00
Xmp.xmpMM.History[2]/stEvt:softwareAgent XmpText   41  Adobe Photoshop Lightroom 5.7.1 (Windows)
Xmp.xmpMM.History[2]/stEvt:changed XmpText    1  /
Xmp.xmpMM.DerivedFrom           XmpText    0  type="Struct"
Xmp.xmpMM.DerivedFrom/stRef:documentID XmpText   32  51711369AC381F3A72B5B60DDF7F26A1
```

Xmp.xmpMM.DerivedFrom/stRef:originalDocumentID	XmpText	32	51711369AC381F3A72B5B60DDF7F26A1
Xmp.xmpRights.Marked	XmpText	4	True
Xmp.xmpRights.WebStatement	XmpText	48	http://creativecommons.org/license/by-nc-sa/3.0/
Xmp.xmpRights.UsageTerms	LangAlt	1	lang="x-default" Creative Commons Attribution-NonC ommerical-ShareAlike
Xmp.crs.HasCrop	XmpText	5	False
Xmp.tiff.Software	XmpText	13	digiKam-4.9.0
Xmp.tiff.DateTime	XmpText	19	2015-05-13T17:12:43
Xmp.tiff.ImageDescription	LangAlt	1	lang="x-default" Pots Caption
Xmp.exif.DateTimeOriginal	XmpText	19	2015:05:13 17:12:43
Xmp.exif.UserComment	LangAlt	1	lang="x-default" Pots Caption
Xmp.digiKam.PickLabel	XmpText	1	0
Xmp.digiKam.ColorLabel	XmpText	1	0
Xmp.digiKam.TagsList	XmpSeq	1	pots
Xmp.digiKam.CaptionsAuthorNames	LangAlt	1	lang="x-default"
Xmp.digiKam.CaptionsDateTimeStamps	LangAlt	1	lang="x-default" 2015-05-15T17:20:37
Xmp.MicrosoftPhoto.Rating	XmpText	2	99
Xmp.MicrosoftPhoto.LastKeywordXMP	XmpBag	1	pots
Xmp.iptc.CreatorContactInfo	XmpText	0	type="Struct "
Xmp.iptc.CreatorContactInfo/Iptc4xmpCore:CiEmailWork	XmpText	26	ericbinaryworld@gmail.com
Xmp.iptc.CreatorContactInfo/Iptc4xmpCore:CiAdrCtry	XmpText	3	USA
Xmp.iptc.CreatorContactInfo/Iptc4xmpCore:CiAdrCity	XmpText	8	Elkridge
Xmp.iptc.CreatorContactInfo/Iptc4xmpCore:CiAdrRegion	XmpText	2	MD
Xmp.iptc.CreatorContactInfo/Iptc4xmpCore:CiUrlWork	XmpText	31	http://www.ericbinaryworld.com
Xmp.mwg-rs.Regions	XmpText	0	type="Struct "
Xmp.mwg-rs.Regions/mwg-rs:RegionList	XmpBag	0	
Xmp.MP.RegionInfo	XmpText	0	type="Struct "
Xmp.MP.RegionInfo/MPRI:Regions	XmpBag	0	
Xmp.lr.hierarchicalSubject	XmpBag	1	pots

#12 - 15 May 2015 23:19 - Alan Pater

JPEG:

<https://www.dropbox.com/s/l8rw3zpjgf599f7/Pots.jpg?dl=0>

That JPEG file was created from the DNG using RawTherapee?

#13 - 15 May 2015 23:49 - Eric Mesa

Alan Pater wrote:

JPEG:

<https://www.dropbox.com/s/l8rw3zpjgf599f7/Pots.jpg?dl=0>

That JPEG file was created from the DNG using RawTherapee?

Correct!

#14 - 15 May 2015 23:49 - Eric Mesa

Alan Pater wrote:

JPEG:

<https://www.dropbox.com/s/l8rw3zpjgf599f7/Pots.jpg?dl=0>

That JPEG file was created from the DNG using RawTherapee?

The Lightroom tags are from when I made the DNG

#15 - 15 May 2015 23:56 - Alan Pater

Eric Mesa wrote:

Alan Pater wrote:

JPEG:

<https://www.dropbox.com/s/l8rw3zpjgf599f7/Pots.jpg?dl=0>

That JPEG file was created from the DNG using RawTherapee?

Correct!

We might be looking at: <https://code.google.com/p/rawtherapee/issues/detail?id=2323>

#16 - 15 May 2015 23:58 - Eric Mesa

Eric Mesa wrote:

Alan Pater wrote:

JPEG:

<https://www.dropbox.com/s/l8rw3zpjgf599f7/Pots.jpg?dl=0>

That JPEG file was created from the DNG using RawTherapee?

The Lightroom tags are from when I made the DNG

What's interesting is that since LR made the first tags, Digikam/Exiv2 put their tags in there too.

#17 - 16 May 2015 00:03 - Alan Pater

Eric Mesa wrote:

The Lightroom tags are from when I made the DNG

What's interesting is that since LR made the first tags, Digikam/Exiv2 put their tags in there too.

Yes, Digikam tries to be as interoperable as possible, mapping a wide variety of proprietary tags as well as following MWG guidelines. Exiv2 just provides the background functions to allow that to happen.

#18 - 16 May 2015 00:20 - Eric Mesa

Alan Pater wrote:

Eric Mesa wrote:

Alan Pater wrote:

JPEG:

<https://www.dropbox.com/s/l8rw3zpjgf599f7/Pots.jpg?dl=0>

That JPEG file was created from the DNG using RawTherapee?

Correct!

We might be looking at: <https://code.google.com/p/rawtherapee/issues/detail?id=2323>

Certainly looks similar. I'll mention it on their bug forums.

#19 - 14 Jan 2016 10:20 - Robin Mills

- Assignee set to Robin Mills
- % Done changed from 0 to 20
- Estimated time set to 5.00 h

I'm not sure what this is about.

[#4184](#) I've submitted a change to src/cr2image.cpp to support options -pS, -pR, -pX, -pC which print the **S**imple Structure, **R**ecursive Structure, **X**MP and ICC **C**olor Profile, so now we can get the RAW XMP in the file:

#23 - 14 Jan 2016 15:43 - Robin Mills

- Estimated time changed from 5.00 h to 10.00 h

Ah, right! Gaaazzunk. The Penny has Dropped!!! Ching.

There are two different file parsers in exiv2. The ones written by Andreas and Brad and the ones I wrote which handle -pX, -pR, -pS, -pC. Mine know about XMLPacket and I suspect that Andreas/Brad parsers don't. However Pots.jpg has revealed an interesting bug in -pX which I'll fix this afternoon. No problem.

Then I'll have a sniff into the Andreas/Brad parser. I'm sure it's broken in the same way which is:

a) XMP is normally stored in JPEGs in an APP1 segment

```
$ exiv2 -pS Stonehenge.jpg
STRUCTURE OF JPEG FILE: Stonehenge.jpg
address | marker | length | data
  2 | 0xd8 SOI | 0
  4 | 0xe1 APP1 | 15288 | Exif..II*.....
15294 | 0xe1 APP1 | 2610 | http://ns.adobe.com/xap/1.0/.<?x
17906 | 0xed APP13 | 96 | Photoshop 3.0.8BIM.....'.....
18004 | 0xe2 APP2 | 4094 | MPF.II*.....0100.....
22100 | 0xdb DQT | 132
22234 | 0xc0 SOF0 | 17
22253 | 0xc4 DHT | 418
22673 | 0xda SOS | 12
$
```

XMP is normally stored in an XMLPacket in a Tiff:

```
$ exiv2 -pS Reagan.tiff
STRUCTURE OF TIFF FILE (MM): Reagan.tiff
address | tag | type | count | offset | value
...
8619248 | 0x02bc XMLPacket | BYTE | 3686 | 8621334 | <x:xmpmeta xmlns:x="adobe:ns:m
et ...
8619260 | 0x83bb IPTCNAA | UNDEFINED | 925 | 8620408 | ...
8619272 | 0x8769 ExifTag | LONG | 1 | 8 | 8
8619284 | 0x8773 InterColorProfile | UNDEFINED | 3144 | 8625020 | ...
...
END Reagan.tiff
$
```

Pots.jpg doesn't have an APP1/XMP packet:

```
$ exiv2 -pS ~/Downloads/Pots.jpg
STRUCTURE OF JPEG FILE: /Users/rmills/Downloads/Pots.jpg
address | marker | length | data
  2 | 0xd8 SOI | 0
  4 | 0xe1 APP1 | 8548 | Exif..II*.....:.....
 8554 | 0xe2 APP2 | 25588 | ICC_PROFILE.....c.lcms.0..mnrRG
34144 | 0xdb DQT | 67
34213 | 0xdb DQT | 67
34282 | 0xc0 SOF0 | 17
34301 | 0xc4 DHT | 29
34332 | 0xc4 DHT | 85
34419 | 0xc4 DHT | 28
34449 | 0xc4 DHT | 72
34523 | 0xda SOS | 12
$
```

It does however have an XMLPacket embedded in the exif data (which is a TIFF formatted data-structure).

```
$ exiv2 -pR ~/Downloads/Pots.jpg
STRUCTURE OF JPEG FILE: /Users/rmills/Downloads/Pots.jpg
address | marker | length | data
  2 | 0xd8 SOI | 0
  4 | 0xe1 APP1 | 8548 | Exif..II*.....:.....
STRUCTURE OF TIFF FILE (II): MemIo
address | tag | type | count | offset | value
  10 | 0x0100 ImageWidth | LONG | 1 | 3898 | 3898
  22 | 0x0101 ImageLength | LONG | 1 | 2594 | 2594
....
 214 | 0x014a SubIFDs | LONG | 5 | 7972 | 7992 8022 8028 8034 8040
 226 | 0x02bc XMLPacket | BYTE | 7246 | 468 | <?xpacket begin="..." id="W5
MOMp ...
```



```

.....
274 | 0x83bb IPTCNAA | LONG | 25 | 7802 | 5898524 1193614083 4260380 1
734960135 1835092841 ...
286 | 0x8769 ExifTag | LONG | 1 | 8046 | 8046
STRUCTURE OF TIFF FILE (II): MemIo
address | tag | type | count | offset | value
8048 | 0x829a ExposureTime | RATIONAL | 1 | 8352 | 8352/0
.....
8336 | 0xa434 LensModel | ASCII | 13 | 8526 | EF50mm f/1.8
END MemIo
298 | 0x9211 ImageNumber | LONG | 1 | 48 | 48
.....
346 | 0xc65d RawDataUniqueID | BYTE | 16 | 7956 | .....6.7
END MemIo
8554 | 0xe2 APP2 | 25588 | ICC_PROFILE.....c.lcms.0..mnrRG
.....
$

```

I suspect this is violation of the XMP/Embedding Specification <http://www.adobe.com/content/dam/Adobe/en/devnet/xmp/pdfs/XMPSpecificationPart3.pdf> Like many Adobe Specs, the document is unreadably boring, however it will be 100% accurate and comprehensive. I'm going to:

- 1) Take the dog for a walk
- 2) Look at <http://www.adobe.com/content/dam/Adobe/en/devnet/xmp/pdfs/XMPSpecificationPart3.pdf>
- 3) Fix my parser to handle Pots.jpg
- 4) Have a read into the Andreas/Brad code. I'm not promising to change that code, however I will investigate.

This team work stuff works. Thank you for the clarification.

#24 - 14 Jan 2016 17:02 - Robin Mills

- File XMPSpecPart3Page75.png added
- % Done changed from 20 to 100
- Estimated time changed from 10.00 h to 3.00 h

Eureka! This is a clear violation of <http://www.adobe.com/content/dam/Adobe/en/devnet/xmp/pdfs/XMPSpecificationPart3.pdf> On page 75 it says:

4.3.1 Metadata storage in JPEG files

JPEG files, including Exif JPEG, use a mixture of native marker segments, TIFF tags, and Photoshop image resources.

Metadata can be found in 3 APPn marker segments:

Marker	Signature, including NULLs	Usage
APP1	"Exif\0\0"	TIFF and Exif (2 NULLs)
APP1	"http://ns.adobe.com/xap/1.0\0"	XMP
APP13	"Photoshop 3.0\0"	Photoshop image resources

The TIFF in the **Exif APP1** marker segment *should not contain tag 700 (XMP), tag 33723 (IPTC), or tag 34377 (Photoshop image resources)*. There should be only one copy of the IPTC in a JPEG file, in Photoshop image resource 1028 (ignoring the possible Mac OS ANPA 10000 resource).

Conclusion:

I don't intend to do further work on this. Somebody will have to swim upstream to determine who wrote this illegal file. I don't believe it was written by Exiv2.

4.3.1 Metadata storage in JPEG files

JPEG files, including Exif JPEG, use a mixture of native marker segments, TIFF tags, and Photoshop image resources. Metadata can be found in 3 APPn marker segments:

Table 47 — APPn marker segments containing XMP metadata

Marker	Signature, including NULLs	Usage
APP1	"Exif0\0"	TIFF and Exif (2 NULLs)
APP1	"http://ns.adobe.com/xap/1.0\0"	XMP
APP13	"Photoshop 3.0\0"	Photoshop image resources

The TIFF in the Exif APP1 marker segment should not contain tag 700 (XMP), tag 33723 (IPTC), or tag 34377 (Photoshop image resources). There should be only one copy of the IPTC in a JPEG file, in Photoshop image resource 1028 (ignoring the possible Mac OS ANPA 10000 resource).

#25 - 15 Jan 2016 02:35 - Eric Mesa

"I don't believe it was written by Exiv2."

Hey,

I'm the OP and I was following your progress today. I created the file with DigiKam and they use Exiv2 for their metadata. Could they be using Exiv2 incorrectly or something?

#26 - 15 Jan 2016 05:39 - Robin Mills

- Status changed from Assigned to Closed

OP? Old Person?

This bug report is in-penetrable. What does this mean (second sentence of your issue report):

Processing is done and a JPEG is produced:

I'm closing this. Please discuss this with DigiKam. I can only deal with issues which can be demonstrated/reproduced with the sample applications distributed with our code.

#27 - 15 Jan 2016 08:59 - Robin Mills

- Status changed from Closed to Assigned

- % Done changed from 100 to 30

- Estimated time changed from 3.00 h to 10.00 h

I'm reopening this issue. We need to painstakingly investigate what has happened here. We can't guess.

I apologise for being so cross. Jumping to the conclusion that Exiv2 is 100% to blame is not helpful. It would also be constructive for you to thank me for the effort I have invested into your issue.

Let's go back to the start. I've download the original CR2 and JPEG files as MG.CR2 and MG.jpg. There is an XMLPacket in the CR2.

```
$ exiv2 -pX MG.CR2 | xmllint --pretty 1 -
<?xml version="1.0"?>
<?xpacket begin="[" id="W5M0MpCehiHzreSzNTczkc9d"?>
<x:xmpmeta xmlns:x="adobe:ns:meta/" x:xmptk="XMP Core 4.4.0-Exiv2">
  <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    <rdf:Description xmlns:tiff="http://ns.adobe.com/tiff/1.0/" ...>
      <tiff:ImageDescription>
        <rdf:Alt>
          <rdf:li xml:lang="x-default">Leaves and Scarlett</rdf:li>
        </rdf:Alt>
      </tiff:ImageDescription>
      <exif:UserComment>
        <rdf:Alt>
          <rdf:li xml:lang="x-default">Leaves and Scarlett</rdf:li>
        </rdf:Alt>
      </exif:UserComment>
      <digikam:TagsList>
        <rdf:Seq>
```

```

    <rdf:li>Scarlett</rdf:li>
  </rdf:Seq>
</digiKam:TagsList>
<digiKam:CaptionsAuthorNames>
  <rdf:Alt>
    <rdf:li xml:lang="x-default"/>
  </rdf:Alt>
</digiKam:CaptionsAuthorNames>
<digiKam:CaptionsDateTimeStamps>
  <rdf:Alt>
    <rdf:li xml:lang="x-default">2015-05-13T20:33:02</rdf:li>
  </rdf:Alt>
</digiKam:CaptionsDateTimeStamps>
<mwg-rs:Regions rdf:parseType="Resource">
  <mwg-rs:RegionList>
    <rdf:Bag/>
  </mwg-rs:RegionList>
</mwg-rs:Regions>
<MP:RegionInfo rdf:parseType="Resource">
  <MPRI:Regions>
    <rdf:Bag/>
  </MPRI:Regions>
</MP:RegionInfo>
<MicrosoftPhoto:LastKeywordXMP>
  <rdf:Bag>
    <rdf:li>Scarlett</rdf:li>
  </rdf:Bag>
</MicrosoftPhoto:LastKeywordXMP>
<lr:hierarchicalSubject>
  <rdf:Bag>
    <rdf:li>Scarlett</rdf:li>
  </rdf:Bag>
</lr:hierarchicalSubject>
<dc:title>
  <rdf:Alt>
    <rdf:li xml:lang="x-default">Scarlett Playing with Leaves</rdf:li>
  </rdf:Alt>
</dc:title>
<dc:description>
  <rdf:Alt>
    <rdf:li xml:lang="x-default">Leaves and Scarlett</rdf:li>
  </rdf:Alt>
</dc:description>
<dc:subject>
  <rdf:Bag>
    <rdf:li>Scarlett</rdf:li>
  </rdf:Bag>
</dc:subject>
</rdf:Description>
</rdf:RDF>
</x:xmpmeta>
<?xpacket end="w"?>
$

```

And this is identified by exiv2:

```

$ exiv2 -px MG.CR2
Xmp.tiff.Software          XmpText    13  digiKam-4.9.0
Xmp.tiff.DateTime         XmpText    19  2015-05-13T17:08:52
Xmp.tiff.ImageDescription  LangAlt     1  lang="x-default" Leaves and Scarlett
Xmp.xmp.CreatorTool       XmpText    13  digiKam-4.9.0
Xmp.xmp.CreateDate        XmpText    19  2015-05-13T17:08:52
Xmp.xmp.MetadataDate      XmpText    19  2015-05-13T17:08:52
Xmp.xmp.ModifyDate        XmpText    19  2015-05-13T17:08:52
Xmp.exif.DateTimeOriginal XmpText    19  2015:05:13 17:08:52
Xmp.exif.UserComment      LangAlt     1  lang="x-default" Leaves and Scarlett
Xmp.photoshop.DateCreated XmpText    19  2015-05-13T17:08:52
Xmp.photoshop.Urgency     XmpText    1  0
Xmp.digiKam.PickLabel     XmpText    1  0
Xmp.digiKam.ColorLabel    XmpText    1  0
Xmp.digiKam.TagsList      XmpSeq      1  Scarlett
Xmp.digiKam.CaptionsAuthorNames LangAlt     1  lang="x-default"
Xmp.digiKam.CaptionsDateTimeStamps LangAlt     1  lang="x-default" 2015-05-13T20:33:02
Xmp.mwg-rs.Regions        XmpText    0  type="Struct"
Xmp.mwg-rs.Regions/mwg-rs:RegionList XmpBag      0

```

```

Xmp.MP.RegionInfo                XmpText    0  type="Struct"
Xmp.MP.RegionInfo/MPRI:Regions    XmpBag     0
Xmp.MicrosoftPhoto.LastKeywordXMP XmpBag     1  Scarlett
Xmp.lr.hierarchicalSubject        XmpBag     1  Scarlett
Xmp.dc.title                      LangAlt    1  lang="x-default" Scarlett Playing with Leaves
Xmp.dc.description                LangAlt    1  lang="x-default" Leaves and Scarlett
Xmp.dc.subject                    XmpBag     1  Scarlett
1865 rmills@rmillsmbp:~/gnu/exiv2/trunk $

```

This is not the original file from the Camera. It has already been modified by DigiKam - as you can see in the Xmp.xmp.CreatorTool.

Next we have the JPEG. Which appears to have NO XMP data at all:

```

$ exiv2 -px MG.JPG
Warning: Ignoring IPTC information encoded in the Exif data.
Warning: Ignoring XMP information encoded in the Exif data.
$

```

It has no APP1/XMP segment as required by the Adobe Spec:

```

$ exiv2 -pS MG.jpg
STRUCTURE OF JPEG FILE: MG.jpg
address | marker | length | data
----- | - | - | -
      2 | 0xd8 SOI |      0 |
      4 | 0xe1 APP1 |  32716 | Exif..II*.....
 32722 | 0xe2 APP2 |  25588 | ICC_PROFILE....c.lcms.0..mnrRG
 58312 | 0xdb DQT |      67 |
 58381 | 0xdb DQT |      67 |
 58450 | 0xc0 SOF0 |      17 |
 58469 | 0xc4 DHT |      29 |
 58500 | 0xc4 DHT |      77 |
 58579 | 0xc4 DHT |      28 |
 58609 | 0xc4 DHT |      70 |
 58681 | 0xda SOS |      12 |
$

```

It has no XMLPacket (as demonstrated by Pots.jpg).

```

$ exiv2 -pR MG.jpg | head -2 ; exiv2 -pR MG.jpg | grep -i xmp
STRUCTURE OF JPEG FILE: MG.jpg
address | marker | length | data
$

```

Conclusion

I don't know how the CR2 was converted into a JPEG. To my knowledge, Exiv2 does not do image format conversion. The XMP has been lost in that conversion. I recommend that you discuss this with DigiKam as I don't believe we performed that conversion.

The discoveries about Pots.jpg are interesting, however I don't understand the genesis or relevance of that file to the discussion.

Please understand that it's hard for me to deal with bugs that occur in applications that use Exiv2. It's easy for DigiKam/Gimp/Geeqie/gThumb and others to say "Metadata issue, must be Exiv2". It's seldom Exiv2's issue. However I will investigate when the issue is condensed and reproducible using our sample applications. Sure that's hard work to isolate and identify the issue. I can't undertake that effort without a lot more input including version and platform. And please understand that I don't use DigiKam/Gimp... so it takes a lot of effort for me to reproduce the application behaviour. It only makes sense to ask the engineers at DigiKam/Gimp... to do that work. I will help you, as I have above, to give you the analysis to present your case to DigiKam/Gimp... The essential information is here. Let me summarise:

- 1) You have a .CR2 which contains metadata. That metadata is correctly identified by Exiv2. *snipp from above*
- 2) You have saved the image as a JPEG. The file no longer has XMP metadata *snipp from above*

I will investigate the meaning of the messages:

```

Warning: Ignoring IPTC information encoded in the Exif data.
Warning: Ignoring XMP information encoded in the Exif data.

```

I didn't write the code that generates those messages, however I'll step the code in the debugger and discover what this is about.

#28 - 15 Jan 2016 09:33 - Robin Mills

- % Done changed from 30 to 40

#29 - 15 Jan 2016 10:06 - Robin Mills

- % Done changed from 40 to 50

I've thought (over breakfast) of how Exiv2 could be responsible for the loss of metadata. I won't explain my thoughts in detail because I have explored and **eliminated** my idea. To convert an image to another format, I believe DigiKam has to:

- 1) Create an empty "container" of the required format
- 2) Copy the metadata from the source to the destination
- 3) Copy the image

We have an empty JPG in our test suite and the sample application **metacopy** can perform the metadata copy. I've reproduced steps 1 and 2 as follows:

```
$ cp test/data/exiv2-empty.jpg .
$ bin/metacopy -a MG.CR2 exiv2-empty.jpg
Warning: Exif tag Exif.Photo.MakerNote not encoded
Warning: Exif tag Exif.Canon.0x4002 not encoded
Warning: Exif tag Exif.Canon.0x4005 not encoded
$ exiv2 -px exiv2-empty.jpg
Xmp.tiff.Software           XmpText      13  digiKam-4.9.0
Xmp.tiff.DateTime          XmpText      19  2015-05-13T17:08:52
Xmp.tiff.ImageDescription  LangAlt       1  lang="x-default" Leaves and Scarlett
Xmp.xmp.CreatorTool       XmpText      13  digiKam-4.9.0
Xmp.xmp.CreateDate        XmpText      19  2015-05-13T17:08:52
Xmp.xmp.MetadataDate      XmpText      19  2015-05-13T17:08:52
Xmp.xmp.ModifyDate        XmpText      19  2015-05-13T17:08:52
Xmp.exif.DateTimeOriginal XmpText      19  2015:05:13 17:08:52
Xmp.exif.UserComment      LangAlt       1  lang="x-default" Leaves and Scarlett
Xmp.photoshop.DateCreated XmpText      19  2015-05-13T17:08:52
Xmp.photoshop.Urgency     XmpText       1  0
Xmp.digiKam.PickLabel     XmpText       1  0
Xmp.digiKam.ColorLabel    XmpText       1  0
Xmp.digiKam.TagsList      XmpSeq        1  Scarlett
Xmp.digiKam.CaptionsDateTimes LangAlt       1  lang="x-default" 2015-05-13T20:33:02
Xmp.mwg-rs.Regions        XmpText       0  type="Struct"
Xmp.mwg-rs.Regions/mwg-rs:RegionList XmpBag        0
Xmp.MP.RegionInfo         XmpText       0  type="Struct"
Xmp.MP.RegionInfo/MPRI:Regions XmpBag        0
Xmp.MicrosoftPhoto.LastKeywordXMP XmpBag        1  Scarlett
Xmp.lr.hierarchicalSubject XmpBag        1  Scarlett
Xmp.dc.title              LangAlt       1  lang="x-default" Scarlett Playing with Leaves
Xmp.dc.description        LangAlt       1  lang="x-default" Leaves and Scarlett
Xmp.dc.subject            XmpBag        1  Scarlett
$
```

The new file has an APP1/xmp segment and does not have a XMLPacket.

```
$ exiv2 -pS empty.jpg
STRUCTURE OF JPEG FILE: empty.jpg
address | marker | length | data
-----|-----|-----|-----
    2 | 0xd8 SOI |      0 |
    4 | 0xe0 APP0 |     16 | JFIF.....H.H....
   22 | 0xe1 APP1 |   2386 | http://ns.adobe.com/xap/1.0/.<?x
  2410 | 0xdb DQT |     67 |
  2479 | 0xdb DQT |     67 |
  2548 | 0xc0 SOF0 |     17 |
  2567 | 0xc4 DHT |     28 |
  2597 | 0xc4 DHT |     60 |
  2659 | 0xc4 DHT |     26 |
  2687 | 0xc4 DHT |     37 |
  2726 | 0xda SOS |     12 |
$ exiv2 -pR empty.jpg | grep -i XMP
$
```

This seems correct to me. I don't know how DigiKam added the metadata during the conversion from CR2 to JPEG - however it doesn't appear to use our recommended code provided in samples/metacopy.cpp as MG.jpg does not have an APP1/XMP data segment:

```
1885 rmills@rmillsmbp:~/gnu/exiv2/trunk $ exiv2 -pS MG.jpg
STRUCTURE OF JPEG FILE: MG.jpg
address | marker | length | data
-----|-----|-----|-----
    2 | 0xd8 SOI |      0 |
    4 | 0xe1 APP1 |  32716 | Exif..II*.....:.....
  32722 | 0xe2 APP2 |  25588 | ICC_PROFILE.....c.lcms.0..mnrRG
  58312 | 0xdb DQT |     67 |
  58381 | 0xdb DQT |     67 |
  58450 | 0xc0 SOF0 |     17 |
  58469 | 0xc4 DHT |     29 |
  58500 | 0xc4 DHT |     77 |
```

58579		0xc4 DHT		28
58609		0xc4 DHT		70
58681		0xda SOS		12

It does however have an ICC profile. I'm adding ICC support in v0.26 and when that is finished, metacopy will be updated appropriately. Until v0.26 ships, Exiv2 does not provide support for ICC profiles. So there is strong evidence that DigiKam has used some other software to do the image conversion and lost the XMP in the process.

#30 - 15 Jan 2016 12:30 - Alan Pater

My understanding is that RawTherapee was used to convert the raw files to jpeg. And that RawTherapee wrongly places the XMP data in XMLPacket. RawTherapee needs to stop doing that. <https://github.com/Beep6581/RawTherapee/issues/2307>

The question for exiv2 is if we can get inside that XMLPacket and return the XMP values to the user.

#31 - 15 Jan 2016 12:38 - Alan Pater

- Tracker changed from Bug to Feature

- Subject changed from Unable to read XMP from CR2 raw file when stored in XMLPacket to Read XMP values from CR2 raw file when stored in XMLPacket

I think this is a feature request rather than a bug in exiv2.

#32 - 15 Jan 2016 13:21 - Robin Mills

For certain this is not a bug in Exiv2. I almost understand this now. I thought DigiKam was involved. It's RawTherapee that is the culprit. He has converted the file to JPEG and preserved the XMLPacket. That's a violation of the Adobe spec. They should put the XMP into the APP1 segment of the JPEG.

And I've realised that I used -pR incorrectly on MG.jpg to grep for xmp. It should be xml

```
$ exiv2 -pR MG.jpg | head -1 ; exiv2 -pR MG.jpg | grep offset ; exiv2 -pR MG.jpg | grep XML
STRUCTURE OF JPEG FILE: MG.jpg
  address | tag | type | count | offset | value
  202 | 0x02bc XMLPacket | BYTE | 4868 | 344 | <?xpacket begin="..." id="W5M0Mp ...
$
```

This is the behaviour observed in Pots.jpg. We're on the same page now. Phew!

Let's discuss the following proposal:

I can easily get -pX to read the XMLPacket. I haven't investigated the Andreas/Brad file parsers, however I think it's probably easy.

However, if we rescue the situation, there will be no pressure on RawTherapee to fix their code. We are blessing a violation of the Adobe spec. You've got a tough sell to persuade me to do that.

We could add an option to exiv2 --fixXMLPacket to repair the JPEG. We have a precedent for this when I agreed to implement -dl to fix files with multiple PhotoShop APP13 segments. #922 I still haven't dealt with -dl yet and maybe we can implement both as the single option --fixJPEG. I don't believe that is blessing a spec violation - quite the opposite. We are providing a utility to enforce the specification. The user has to run the command, we shouldn't do this automatically.

If we decide to adopt this solution, I don't need to look at the Andreas/Brad file parsers and they will continue to issue their warnings. This is good. Andreas and Brad have done a great job and I will only modify their code if I find a bug.

I don't think we should be in hurry to implement this. Perhaps some pressure can be applied to RawTherapee. On this evidence, they don't appear to be using Exiv2. I'm willing to work with them to integrate libexiv2 into their product. However I don't have time to work with them until v0.26 is code-complete in April.

Thoughts?

#33 - 15 Jan 2016 14:28 - Alan Pater

Eric, please work with the RawTherapee folks to fix this on their end.

Robin, for this feature to be useful to users on DigiKam and other existing applications, it needs to work with the default exiv2 command line. "exiv2 pots.jpg" should display the XMP values hidden in XMLPacket.

Fixing data errors in broken image files is another can of worms. If exiv2 can provide a way to read the data, other utilities can use that to fix broken files. I don't think exiv2 needs that feature internally when it can be provided by an external program.

#34 - 15 Jan 2016 15:18 - Robin Mills

One thing is 100% certain. There is no bug in Exiv2. The behaviour of our code is correct.

If we read that file and silently forgive the error, RawTherapee will do nothing about their bug.

There is a second spec violation. IPTC should be stored in the APP13/Photoshop segment of a JPEG. IPTC data should not be stored in an IPTCNAA tag within the Exif data:

```
$ exiv2 -pR MG.jpg | head -1 ; exiv2 -pR MG.jpg | grep address ; exiv2 -pR MG.jpg | grep IPTC
STRUCTURE OF JPEG FILE: MG.jpg
  address | tag | type | count | offset | value
    214 | 0x83bb IPTCNAA | LONG | 33 | 5212 | 5898524 1193614083 4260380 1734960135 1835092841 ...
$
```

The Andreas/Brad file parser correctly issues warning and ignores those tags.

```
Warning: Ignoring IPTC information encoded in the Exif data.
Warning: Ignoring XMP information encoded in the Exif data.
```

Here's a new proposal:

- 1) *We close this issue.*
- 2) *We open a new feature request and set the Target version to 1.0*
This new issue can skip most of the detail discussed here (however the new issue should be linked to this).

I've asked Andreas to review all 1.0 issues before we ship v0.26. We can add Phil as a watcher on the new issue. Phil is really smart and will say something interesting and clever about this.

#35 - 15 Jan 2016 16:52 - Robin Mills

- *Tracker changed from Feature to Bug*
- *Status changed from Assigned to Closed*
- *% Done changed from 50 to 100*

I'm going to close this. It's not a bug.

I'll open a new issue to consider reading the illegal XMP and IPTC data from the Tiff-Encoded Exif. The target for that issue will be 1.0. So it will be reviewed as part of the v0.26 release process and a decision made about the best course of action.

#36 - 16 Jan 2016 06:49 - Robin Mills

- *Assignee changed from Robin Mills to Alan Pater*

I'm assigned this to Alan. Alan's the guy who really understood this issue and dealt with RawTherapee. Great Job, Alan.

Files

XMPSpecPart3Page75.png	106 KB	14 Jan 2016	Robin Mills
------------------------	--------	-------------	-------------