

Exiv2 - Feature #937

Support Darwin Core (DwC) XMP metadata

04 Dec 2013 12:04 - Jim Nelson

Status:	Closed	Start date:	04 Dec 2013
Priority:	Normal	Due date:	
Assignee:	Alan Pater	% Done:	100%
Category:	xmp	Estimated time:	0.00 hour
Target version:	0.25		
Description			
There has been a request from the GIMP project to support Darwin Core (DwC) metadata. This was discussed on the Exiv2 forums at http://dev.exiv2.org/boards/3/topics/1675 but I couldn't find a ticket for it.			
More information about Darwin Core can be found at http://rs.tdwg.org/dwc/			
Related issues:			
Related to Exiv2 - Feature #941: Upgrade xmpsdk source to Adobe's current ver...		Closed	27 Dec 2013

Associated revisions

Revision 3212 - 23 Dec 2013 14:58 - Robin Mills

Issue: #937. Thanks to Alan and Jim for raising the issue. Thanks to Alan for the patch and test file.

Revision 3266 - 25 Jun 2014 18:04 - Robin Mills

See #937 (DwC Darwin Core Support).

Revision 3267 - 26 Jun 2014 02:34 - Robin Mills

#937. Thanks to Alan for the patch code and data file.

Revision 3658 - 04 Apr 2015 17:23 - Alan Pater

#937 Darwin Core 2015-03-19 schema update, plus doc template for same

History

#1 - 08 Dec 2013 23:04 - Alan Pater

- File *DwCproperties.cpp* added

- File *DwC-SampleImage.jpg* added

I slapped together *DwCproperties.cpp* but it is likely missing some important details, specifically I am unclear on how things like *xmpText*, *xmpExternal*, "Integer", *signedLong*, and *xmpInternal* are determined.

I've also attached a sample image with DwC metadata.

#2 - 09 Dec 2013 22:54 - Robin Mills

- File *DwC2.txt* added

- File *DwC.patch* added

Guys

I have a confession to make. I didn't write Exiv2 and don't exactly know how it works. I have to read the code to figure it out (or hope Andreas will help if/when I'm in trouble). I suspect *xmpText* (and cohorts) relate to the XMP specification which I have not read. However, I have no fear of the unknown and push on with enthusiastic uncertainty and use the time honored software development approach of adding code, then finding out if it's wrong!

I have positive progress to report. I grafted your code into *properties.cpp*, built and ran *exiv2* against your sample image. Good News. It's producing impressive output:

```
566 rmills@rmills-mbp:~/gnu/exiv2/trunk $ bin/.libs/exiv2 -pa DwC2.jpg 2> /dev/null | grep dwc | wc
    173     1015     14841
567 rmills@rmills-mbp:~/gnu/exiv2/trunk $ bin/.libs/exiv2 -pa DwC2.jpg 2> /dev/null | grep dwc
Xmp.dwc.Record          XmpText          0  type="Struct "
```

```

Xmp.dwc.Record/dwc:institutionID      XmpText    25  Charles Darwin Foundation
Xmp.dwc.Record/dwc:collectionID      XmpText    29  urn:lsid:biocol.org:col:34818
Xmp.dwc.Record/dwc:institutionCode    XmpText    3   CDS
Xmp.dwc.Record/dwc:datasetID         XmpText    3   MVZ
Xmp.dwc.Record/dwc:collectionCode    XmpText    7   Mammals
Xmp.dwc.Record/dwc:datasetName       XmpText    25  Grinnell Resurvey Mammals
Xmp.dwc.Record/dwc:ownerInstitutionCode XmpText    3   NPS
...
Xmp.dwc.MeasurementOrFact/dwc:measurementDeterminedDate XmpText    25  2013-01-27T00:00:00-06:00
Xmp.dwc.MeasurementOrFact/dwc:measurementDeterminedBy XmpText    18  Javier de la Torre

```

I assume Grinnell Resurvey Mammals is something to do with Mr Grinnell of Grinnell Glacier, Montana. I've been there (photo below).

I attach a couple of files:

- 1) The output of `exiv2 -pa DwC2.jpg` (I shorted the filename from `DwC-SampleImage.jpg`)
- 2) A Patch (which you can apply to the current trunk [r3209](#))

I've run our test suite (on the Mac) and it passes. We released 0.24 last week - so this will probably be the first submission for 0.25. The patch can be applied to the 0.24 release. As this is a graft to existing code, there are no build changes and I expect this will run on all supported platforms and build environments.

Your code had formatting difficulties with `""` (consecutive double-quotes). I suspect your code was machine generated from Perl and isn't valid C++. I've fixed it to compile. Perhaps you can review, repair my changes and attach your version of the patch. I'll be happy to submit your version to the trunk provided it compiles and passes the test suite. I will also update the test suite with your test image to guarantee DwC support going forward.

Robin
Grinnell.jpg

#3 - 09 Dec 2013 23:14 - Robin Mills

- Status changed from New to Assigned
- Assignee set to Robin Mills
- Target version set to 0.25
- % Done changed from 0 to 50

#4 - 10 Dec 2013 15:58 - Alan Pater

Very cool, thanks a bunch Robin. I'm downloading the trunk and the patch as we speak and will try to figure out where to go from here. My coding skills are pretty basic ...

Just as a note, I ran an older version of `exiv2` tool on the DwC image and, as a comparison, `exiftool`. `exiv2` appears to be outputting the raw XMP data.

```

asp@devel:~$ exiv2 -pa DwC-SampleImage.jpg | grep Grinnel
Warning: Unsupported time format

```

```

Xmp.dwc.Event/dwc:fieldNotes          XmpText    42  notes available in Grinnell-Miller Library
Xmp.dwc.Record/dwc:datasetName        XmpText    25  Grinnell Resurvey Mammals

```

```

asp@devel:~$ exiftool DwC-SampleImage.jpg | grep Grinnel

```

```

Event Field Notes          : notes available in Grinnell-Miller Library
Record Dataset Name       : Grinnell Resurvey Mammals

```

#5 - 10 Dec 2013 16:44 - Robin Mills

Hold your guns there, buster.

I don't believe your code is achieving anything. When I saw all the dwc output, I assumed it came from your code. I've just done a rebuild of the trunk (without your code). The output is the same with/without your code!

Which begs two obvious questions:

- 1) Is the current output from `Exiv2` sufficient for your purposes?
- 2) what is your code intended to achieve?

As far as fixing the `""` (double quotes). In C++, a string is `"....."`. If you want to embed a double quote, you say `\"` within the string. For example `"I am a \"Rockstar\" you know"`. Some of your strings are a mostly incompressible mix of " comma space (e.g. line#168). Perhaps you know the intended string.

Robin

#6 - 10 Dec 2013 16:53 - Alan Pater

Looks that way, doesn't it? Sorry for the bad attempt, I have a lot to learn about this stuff!

I see that exiv2 is capable of displaying raw XMP data, even if it is not cognizant of the underlying schema. What I am hoping for is that we can get it capable of fully understanding the schema and therefor being able to write and edit DwC fields.

I also don't have a good understanding of how exiv2 works. Are there other functions that take care of editing metadata fields and of displaying those fields in a pretty manner?

#7 - 10 Dec 2013 16:57 - Robin Mills

Let me read the code relating to this after work and see what's going on. I'll get back to you.

Robin

#8 - 10 Dec 2013 20:42 - Robin Mills

Guys

I'm going to have to ask Andreas for help with this.

Andreas:

Do you know how the Xmp and properties get connected? We can list the "raw" XMP data, however I can't change it.

```
753 rmills@rmills-mbp:~/gnu/exiv2/gsoc13 $ exiv2 -pa DwC2.jpg | grep -i missing
Warning: Unsupported time format
Xmp.dwc.MeasurementOrFact/dwc:measurementRemarks XmpText 19 tip of tail missing
754 rmills@rmills-mbp:~/gnu/exiv2/gsoc13 $ exiv2 -M"set Xmp.dwc.MearsurementOrFact/dwc:measurementRemarks xmp
Text Missing" DwC2.jpg
-M option 1: Invalid key `Xmp.dwc.MearsurementOrFact/dwc:measurementRemarks'
exiv2: Error parsing -M option arguments
Usage: exiv2 [ options ] [ action ] file ...
```

Manipulate the Exif metadata of images.

```
755 rmills@rmills-mbp:~/gnu/exiv2/gsoc13 $
```

The code that Jim and Alan have provided seems to set up the correct mapping. I've got their code to build, however it has no effect. Can you help us, please?

Guys

Just, so you don't think that I am lazy, I have spend a couple of hours in the debugger without making progress. So I turned to a different tool. In our gsoc13 branch, we have a new feature for 0.25 which enables dumping the structure of a JPG. Currently, this is exercised from samples/exifprint.

```
733 rmills@rmills-mbp:~/gnu/exiv2/gsoc13 $ bin/exifprint --struc DwC2.jpg
STRUCTURE OF FILE:
  offset | marker | size | signature
      2 | 0xd8 SOI |      | 
      4 | 0xe0 APP0 | 16 | JFIF.....`..
     22 | 0xe1 APP1 | 2015 | Exif..II*.....b.....
    2039 | 0xe1 APP1 | 15936 | http://ns.adobe.com/xap/1.0/.<
   17977 | 0xed APP13 | 3164 | Photoshop 3.0.8BIM.....7.
   21143 | 0xee APP14 | 14 | Adobe.d.....
   21159 | 0xdb DQT | 132 | 
   21293 | 0xc0 SOF0 | 17 | 
   21312 | 0xdd DRI | 4 | 
   21318 | 0xc4 DHT | 418 | 
   21738 | 0xda SOS | 12 | 
-----
```

We can then take the numbers for the position of the APP1 xap block, extract it - and push it through xmllint

```
734 rmills@rmills-mbp:~/gnu/exiv2/gsoc13 $ dd skip=$((2039+32-1)) bs=1 count=$((17977-33-2039)) if=DwC2.jpg 2>
/dev/null | head -2 | tail -1 | xmllint --format 2 - 2>/dev/null
<?xml version="1.0"?>
<x:xmpmeta xmlns:x="adobe:ns:meta/" x:xmptk="Toolkit=IDImager;Version=2.4.0.9;">
  <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
    .....
    <dwc:MeasurementOrFact>
      <rdf:Description xmlns:dwc="http://rs.tdwg.org/dwc/index.htm" rdf:about="">
        <dwc:measurementID>1234</dwc:measurementID>
        <dwc:measurementType>tail length</dwc:measurementType>
        <dwc:measurementValue>45</dwc:measurementValue>
        <dwc:measurementAccuracy>0.01</dwc:measurementAccuracy>
```

```
<dwc:measurementUnit>mm</dwc:measurementUnit>
<dwc:measurementDeterminedDate>2013-01-27T00:00:00-06:00</dwc:measurementDeterminedDate>
<dwc:measurementDeterminedBy>Javier de la Torre</dwc:measurementDeterminedBy>
<dwc:measurementMethod>barometric altimeter</dwc:measurementMethod>
<dwc:measurementRemarks>tip of tail missing</dwc:measurementRemarks>
</rdf:Description>
</dwc:MeasurementOrFact>
</rdf:Description>
<rdf:Description xmlns:MicrosoftPhoto="http://ns.microsoft.com/photo/1.0" rdf:about="" MicrosoftPhoto:Rating="75"/>
</rdf:RDF>
</x:xmpmeta>
```

I don't know that this is useful, however it's fun. If the xap block is rewritten externally, the JPG can be reassembled using dd. I discussed this a few months ago on a thread in connection with IPTC data. <http://dev.exiv2.org/boards/3/topics/1608>

Robin

#9 - 10 Dec 2013 20:43 - Robin Mills

- Assignee changed from Robin Mills to Andreas Huggel

- % Done changed from 50 to 30

#10 - 21 Dec 2013 14:29 - Alan Pater

- File dwc.patch added

Hi folks

I finally found some time to dig in and try to create a valid patch. I've cleaned up the formatting a bit, hopefully it is an improvement and not as ugly as the previous version.

Unfortunately, I have not been able to compile the result, my poor computer seems to be dying. I'll give it a try again later and see how it goes.

Apart from cleaning up the formatting on the list of DwC properties, I've added a line to xmp.cpp:
SXMPMeta::RegisterNamespace("http://rs.tdwg.org/dwc/terms/", "dwc");

Hopefully that will be help to get things working.

#11 - 21 Dec 2013 14:44 - Robin Mills

This isn't my lucky days for patches using SmartSVN. It won't accept your patch - and it wouldn't accept one yesterday from another user. Grrrrr software. I've replaced all trunk.dwc with trunk in the file, without success.

Did you create this with the command `svn diff > foo.patch` ?

Robin

#12 - 21 Dec 2013 14:44 - Robin Mills

- Assignee changed from Andreas Huggel to Robin Mills

#13 - 21 Dec 2013 15:03 - Alan Pater

Robin, yes, I used `svn diff`, but I am brand new to `svn` so I likely messed something up. I may not have setup or are using `svn` correctly.

#14 - 21 Dec 2013 15:20 - Robin Mills

Thanks. I'll get it to work. I'll deal with this tomorrow. We're off to dinner at friends in San Francisco this evening.

Robin

#15 - 21 Dec 2013 15:23 - Robin Mills

Hey, you've only changed a couple of files. Can you send me them and we'll be done in a few minutes.

xmp.cpp and properties.cpp

Robin

```
599 rmills@rmills-mbp:~/gnu/exiv2/trunk $ grep trunk dwc2.patch
Index: trunk/src/xmp.cpp
--- trunk/src/xmp.cpp      (revision 3211)
+++ trunk/src/xmp.cpp      (working copy)
```

```
Index: trunk/src/properties.cpp
--- trunk/src/properties.cpp      (revision 3211)
+++ trunk/src/properties.cpp      (working copy)
600 rmills@rmills-mbp:~/gnu/exiv2/trunk $
```

#16 - 21 Dec 2013 15:52 - Alan Pater

- File *xmp.cpp* added
- File *properties.cpp* added

Here you go :-)

#17 - 22 Dec 2013 09:54 - Alan Pater

Well, I convinced my poor computer to build the package, and it seems to be able to modify DwC fields. However, output still appears to be in a raw XMP format:

```
~$ exiv2 -px dwc.jpg | grep Count
Xmp.dwc.Occurrence/dwc:individualCount      XmpText      1  1
```

While I would expect something like:

```
Xmp.DwC.IndividualCount      XmpText      1  1
```

On the other hand, it looks like I can modify the contents of this field

```
~$ exiv2 -M"add Xmp.dwc.Occurrence/dwc:individualCount XmpText 3" dwc.jpg
~$ exiv2 -px dwc.jpg | grep Count
Xmp.dwc.Occurrence/dwc:individualCount      XmpText      1  3
```

So that part is working, without my patch the result of that command is

```
-M option 1: Invalid key `Xmp.dwc.Occurrence/dwc:individualCount'
```

So I have a bit more work to do to figure out how to display DwC tags nicely.

#18 - 23 Dec 2013 15:03 - Robin Mills

Thanks, Alan. I've submitted your code and extended *test/bugfixes-test.sh* to use your test file. [r3212](#). It's clear that although we're making progress, we're not done here. However it's better than before, so it's worth submitting. I expect we'll be talking more about this in 2014. Maybe Andreas will have a little time during the holidays to really fix this for us.

Anyway. Ho Ho Ho. Happy Holidays.

Robin

#19 - 26 Dec 2013 15:23 - Alan Pater

- File *dwc.patch* added

This version cleans up the formatting, removes duplicate tags and adds some doc files.

I tested it on a wider (but still limited) range of fields and did not see any errors.

However, it still does not display things nicely compared to other XMP namespaces.

```
~$ exiv2 -PXkyctl DwC-SampleImage.jpg | grep Date
Xmp.xmp.CreateDate      Create Date      XmpText      24  2008-03-14T20:59:26
.535Z
Xmp.xmp.MetadataDate    Metadata Date    XmpText      29  2013-02-07T21:56:33
.820-06:00
Xmp.xmp.ModifyDate      Modify Date      XmpText      25  2013-01-27T14:02:29
-06:00
Xmp.dwc.Event/dwc:earliestDate    Event/dwc:earliestDate    XmpText      25  2012-09-03T00:00:00
-06:00
Xmp.dwc.Event/dwc:latestDate      Event/dwc:latestDate      XmpText      25  2013-01-27T00:00:00
-06:00
Xmp.dwc.Event/dwc:verbatimEventDate    Event/dwc:verbatimEventDate    XmpText      11  spring 1910
```

Might it have something to do with the fact that the DwC namespace involves nested tags? Would I have to something special to accommodate that

structure?

#20 - 26 Dec 2013 21:45 - Robin Mills

Thanks very much, Alan. I have submitted [r3213](#). I reverted the changes to xmp.cpp and properties.cpp in [r3212](#) and applied your patch. I also added a variant of your command to the test suite (test/bug-fixes.sh 937).

```
560 rmills@rmills-mbp:~/gnu/exiv2/trunk $ exiv2 -q -PXkyctl -g Date test/data/exiv2-bug937.jpg
Xmp.exif.DateTimeDigitized      Date and Time Digitized      XmpText      29 2008-03-14T11:31:48
.098-07:00
Xmp.exif.DateTimeOriginal      Date and Time Original      XmpText      25 2008-03-14T13:59:26
-06:00
Xmp.photoshop.DateCreated      Date Created      XmpText      29 2008-03-14T13:59:26
.054-06:00
Xmp.xmp.MetadataDate          Metadata Date      XmpText      29 2013-02-07T21:56:33
.820-06:00
Xmp.xmp.CreateDate            Create Date      XmpText      24 2008-03-14T20:59:26
.535Z
Xmp.xmp.ModifyDate            Modify Date      XmpText      25 2013-01-27T14:02:29
-06:00
Xmp.dwc.Event/dwc:earliestDate Event/dwc:earliestDate      XmpText      25 2012-09-03T00:00:00
-06:00
Xmp.dwc.Event/dwc:latestDate  Event/dwc:latestDate      XmpText      25 2013-01-27T00:00:00
-06:00
Xmp.dwc.Event/dwc:verbatimEventDate Event/dwc:verbatimEventDate XmpText      11 spring 1910
Xmp.dwc.ResourceRelationship/dwc:relationshipEstablishedDate ResourceRelationship/dwc:relationshipEstablishedDate XmpText      25 2013-01-27T00:00:00-06:00
Xmp.dwc.MeasurementOrFact/dwc:measurementDeterminedDate MeasurementOrFact/dwc:measurementDeterminedDate XmpText      25 2013-01-27T00:00:00-06:00
561 rmills@rmills-mbp:~/gnu/exiv2/trunk $
```

Thanks very much for persevering with this. I especially appreciate that you updated the documentation. I don't recall any of our users providing a documentation update. Thanks.

Does dwc use "a language within a language" with syntax such as Xmp.dwc.Event/dwc:verbatimEventDate? Should xmpsdk be handling this, or do we need an extension to xmpsdk to handle dwc? What do you think?

Robin

#21 - 27 Dec 2013 04:30 - Alan Pater

Robin, thank you for keeping on top of my changes.

I don't think we need to extend the xmpsdk. My reading of the XMP specification (part 1, 6.3.3) indicates that these types of structures are part of the specification. I left out these base structures in my changes, as I don't understand how to implement them and was not sure that they are needed. Then again, if they are part of the DwC namespace, I suppose they need to be implemented. But how?

It looks to me like DwC has 9 base structures, with the individual tags nested under those:

Xmp.dwc.Record	XmpText	0	type="Struct "
Xmp.dwc.Occurrence	XmpText	0	type="Struct "
Xmp.dwc.Event	XmpText	0	type="Struct "
Xmp.dwc.dctermsLocation	XmpText	0	type="Struct "
Xmp.dwc.GeologicalContext	XmpText	0	type="Struct "
Xmp.dwc.Identification	XmpText	0	type="Struct "
Xmp.dwc.Taxon	XmpText	0	type="Struct "
Xmp.dwc.ResourceRelationship	XmpText	0	type="Struct "
Xmp.dwc.MeasurementOrFact	XmpText	0	type="Struct "

#22 - 27 Dec 2013 19:35 - Robin Mills

Alan

I'd like to get all of this fixed for you, however the only way I can do that is to study XMP and the xmpsdk. Would you like to talk 1-to-1 on Skype about this? Neither of us know how to fix this, however together we might make progress. And of course I'm having a break this week from work.

I downloaded the latest version from Adobe today and I can see the code's changed. Time for us to refresh xmpsdk in our code base. However that'll have to wait.

Robin

#23 - 29 Dec 2013 15:59 - Alan Pater

Yeah, looks like the update to the xmpsdk is a bit overdue: <http://dev.exiv2.org/issues/742>

That said, exiv2 does provide for XmpStruct. It just looks like nobody has used it until now, so I don't have an example I can just adapt. I'm not a programmer, so get lost really quickly looking through the source.

#24 - 29 Dec 2013 17:29 - Robin Mills

Alan

I have opened another issue [#941](#) to update the SDK. There has been very little discussion about XMP on the forum and issue reports, so it hasn't demanded attention. I will look at this, however I'm not promising anything. I don't get lost in code, however it usually takes time to get oriented with something on which I haven't worked.

Robin

#25 - 29 Dec 2013 19:40 - Andreas Huggel

Sorry Robin for not responding earlier.

[r2031](#) is a minimal example of the changes necessary to add support for a new XMP schema, the related doc was added with [r2248](#) and [r2252](#). It's really mostly about making the new namespace known to Exiv2. You should be able to work with a schema that Exiv2 doesn't know just as well, except that you'd have to register the namespace first, there is [an example](#) on the homepage which shows how.

We don't have much support for nested tags, there is no built-in way to make these look better (remove the namespace). In fact, the way we deal with XMP namespaces in general is a bit too simplistic.

HTH

#26 - 30 Dec 2013 10:27 - Alan Pater

Hi Andreas

I looked at those revisions (along with others) to get an idea of what changes would be needed, and it works, write support for Dwc XMP fields is working. Dwc fields can be set and modified using exiv2, interchangeably with exiftool. I can use either tool to write and modify the field values, and the other tool has no problems with the changes. This is good! So maybe we don't really need nesting?

Thanks for clarifying that there is no current way to get them looking prettier, Feels good to stop beating my head against the wall!

I've added the class level fields to my working copy, but without any nesting, so it doesn't really do anything that I can tell. I don't know how to implement that, if it is even possible. Perhaps a real programmer could take a look at the following to see how to nest the terms under the class level?

```
extern const XmpPropertyInfo xmpDwcInfo[] = {
    // Material Sample Level Class
    { "MaterialSample",          N_("Material Sample"),          "bag Struct",    xm
pBag,    xmpInternal,
        N_("The category of information pertaining to the physical results
of a sampling (or subsampling) event. In biological collections, the material sample is typically collected,
and either preserved or destructively processed."),
    },
    // Material Sample Level Terms
    { "materialSampleID",       N_("Material Sample ID"),          "Text",         xmpT
ext,    xmpExternal,
        N_("An identifier for the MaterialSample (as opposed to a part
icular digital record of the material sample). In the absence of a persistent global unique identifier, constr
uct one from a combination of identifiers in the record that will most closely make the materialSampleID globa
lly unique.")
    },
    // End of list marker
    { 0, 0, 0, invalidTypeId, xmpInternal, 0 }
};
```

#27 - 01 Jan 2014 15:00 - Alan Pater

I said that it was working interchangeably with exiftool but I take that back after a bit more thought and testing.

It CAN work interchangeably with exiftool IF one specifies the full nested tag: **Xmp.dwc.dctermsLocation/dwc:verbatimElevation**.

But not interchangeably if one specifies the flat tag: **Xmp.dwc.verbatimElevation**.

That is, my patch of exiv2 can write to both tags, but they are separate. Exiftool can only write to full nested tag.

As exiftool is following the Dwc schema, it is correct. My version is incorrect in that it writes to non-nested fields as well.

#28 - 01 Jan 2014 16:33 - Andreas Huggel

My version is incorrect in that it writes to non-nested fields as well.

Only if you instruct it to do so, right? Meaning, it doesn't refuse to add an unknown property to the dwc namespace but I presume it doesn't write the flat one at the same time with the nested property by itself? That is how Exiv2 is designed to work in general. It won't restrict you to only known properties and leaves the control to the user. Exiftool may be more restrictive (by default?), but as long as Exiv2 can deal with the correct nested property, the test should be considered passed. On the other hand, if you write a non-standard tag, then it's up to the other software what that wants to do with it.

#29 - 02 Jan 2014 08:52 - Alan Pater

Well, I wouldn't want to upset the DwC schema by ignoring its chosen hierarchy. ;-)

But seriously, I'd be a bit nervous about allowing non-nested DwC properties to be written. If exiv2 was the only game in town, that would not be an issue, we could just implement a flat schema and everyone would have to follow. But I feel that we should maintain 100% compatibility with exiftool and any other tools that may come up in the future. If exiv2 can write properties that the others can't, we could be forcing some users to write two tags for each desired property, just to maintain compatibility.

At the moment, when I view the generated HTML documentation, users can see that flat properties are available. We could hack the documentation to only show the correctly nested properties, but I suspect that someone will discover the non-nested properties in other ways. Those non-nested properties fall outside of the DwC schema, following the schema is the objective, no?

Or is there a way to completely hide the non-nested properties? Or alias them to nested properties? Seems like that would be extra work compared to implementing a nested structure in the first place.

#30 - 19 Jan 2014 12:14 - Alan Pater

I am running into a bit of a glitch with some DwC fields. If I write several alt-lang fields in a image, Adobe apps can no longer see the data in simple text fields. To test, I wrote the same fields using both my patched version of exiv2 and exiftool. Adobe apps can see all the data of the image written using exiftool, but only the vernacularName field for the exiv2 version. Extracting the XMP data shows the following difference in how exiftool & exiv2 are writing the fields:

```
// DwC written using: exiv2 -m taxon.txt DwC-taxon.jpg
```

```
<?xpacket begin="<U+FEFF>" 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 rdf:about="" xmlns:dwc="http://rs.tdwg.org/dwc/terms/">
    <dwc:Taxon>
      <rdf:Description dwc:class="Vertebrata" dwc:family="Felidae" dwc:genus="Puma" dwc:infraspecificEpithet="concolor" dwc:kingdom="Animalia" dwc:order="Mammalia" dwc:phylum="Chordata" dwc:specificEpithet="concolor" dwc:subgenus="Puma">
        <dwc:vernacularName>
          <rdf:Alt>
            <rdf:li xml:lang='en-US'>Cougar</rdf:li>
            <rdf:li xml:lang='es-ES'>Puma</rdf:li>
            <rdf:li xml:lang='fr-FR'>Puma</rdf:li>
          </rdf:Alt>
        </dwc:vernacularName>
      </rdf:Description>
    </dwc:Taxon>
  </rdf:Description>
</rdf:RDF>
</x:xmpmeta>
<?xpacket end="w"?>
```

```
// DwC written using: exiftool -csv=taxon.csv DwC-taxon.jpg
```

```
<?xpacket begin="<U+FEFF>" id="W5M0MpCehiHzreSzNTczkc9d"?>
<x:xmpmeta xmlns:x="adobe:ns:meta/" x:xmptk="Image::ExifTool 9.27">
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">

  <rdf:Description rdf:about="" xmlns:dwc="http://rs.tdwg.org/dwc/terms/">
    <dwc:Taxon rdf:parseType='Resource'>
      <dwc:class>Vertebrata</dwc:class>
      <dwc:family>Felidae</dwc:family>
      <dwc:genus>Puma</dwc:genus>
      <dwc:infraspecificEpithet>concolor</dwc:infraspecificEpithet>
      <dwc:kingdom>Animalia</dwc:kingdom>
      <dwc:order>Mammalia</dwc:order>
      <dwc:phylum>Chordata</dwc:phylum>
```



```

<dc:specificEpithet>concolor</dc:specificEpithet>
<dc:subgenus>Puma</dc:subgenus>
<dc:vernacularName>
  <rdf:Alt>
    <rdf:li xml:lang='en-US'>Cougar</rdf:li>
    <rdf:li xml:lang='es-ES'>Puma</rdf:li>
    <rdf:li xml:lang='fr-FR'>Puma</rdf:li>
  </rdf:Alt>
</dc:vernacularName>
</dc:Taxon>
</rdf:Description>
</rdf:RDF>
</x:xmpmeta>
<?xpacket end='w'?>

```

Do I need to be writing these fields differently? exiv2 is dumping all the text fields into a second <rdf:Description ... /> structure.

#31 - 20 Jan 2014 06:37 - Phil Harvey

Alan Pater wrote:

Do I need to be writing these fields differently? exiv2 is dumping all the text fields into a second <rdf:Description ... /> structure.

The second description shouldn't be a problem. ExifTool avoids this with the "rdf:parseType='Resource'" attribute, but this isn't the only way to do it.

The problem seems to be that the Exiv2 has only written the vernacularName, and nothing else.

- Phil

#32 - 20 Jan 2014 06:42 - Phil Harvey

Alan Pater wrote:

Do I need to be writing these fields differently? exiv2 is dumping all the text fields into a second <rdf:Description ... /> structure.

The second description shouldn't be a problem. ExifTool avoids this with the "rdf:parseType='Resource'" attribute, but this isn't the only way to do it.

The problem seems to be that the Exiv2 has only written the vernacularName, and nothing else.

- Phil

P.S. I am a bit confused about the problems you seem to be having with the dwc structure. The standard exif schema has similar one-level structures.

#33 - 20 Jan 2014 09:09 - Alan Pater

I think most of my confusion comes from my lack of a solid programming background ...

To sum up that example, it seems that Adobe apps have trouble reading the fields in this structure:

```
<rdf:Description dwc:class="Vertebrata" dwc:family="Felidae" ...>
```

Here's another example of what exiv2 is writing and that Adobe apps don't understand. This example tries to write both DC & DWC fields:

```
<rdf:Description rdf:about="" xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dwc="http://rs.tdwg.org/dwc/terms/" dc:source="Alan Pater">
```

```

  <dc:type>
    <rdf:Bag>
      <rdf:li>
        test image with DC
      </rdf:li>
    </rdf:Bag>
  </dc:type>

```

```

  <dwc:Taxon dwc:kingdom="Animalia" dwc:phylum="Chordata" dwc:class="Vertebrata" dwc:order="Mammalia" dwc:family="Felidae" dwc:genus="Puma" dwc:subgenus="Puma" dwc:specificEpithet="concolor"/>

```

```
</rdf:Description>
```

Shouldn't the two different namespace have separate rdf:Description containers? Something like:

```
<rdf:Description rdf:about="" xmlns:dc="http://purl.org/dc/elements/1.1/">
```

```
  <dc:source>"Alan Pater"</dc:source>
```

```
  <dc:type>
```

```
    <rdf:Bag>
```

```
      <rdf:li>
```

```
        test image with DC
```

```
      </rdf:li>
```

```
    </rdf:Bag>
```

```
  </dc:type>
```

```
</rdf:Description>
```

```
<rdf:Description rdf:about="" xmlns:dwc="http://rs.tdwg.org/dwc/terms/">
```

```
  <dwc:Taxon dwc:kingdom="Animalia" dwc:phylum="Chordata" dwc:class="Vertebrata" dwc:order="Mammalia" dwc:family="Felidae" dwc:genus="Puma" dwc:subgenus="Puma" dwc:specificEpithet="concolor"/>
```

```
</rdf:Description>
```

#34 - 21 Jan 2014 05:40 - Phil Harvey

Ah, yes. I missed that the structure elements were added as attributes of the Description. You can't mix attributes and elements like this. From the XMP 2012 specification Part1, section 7.9.2.4: "All fields of a structure shall be written in the same manner, either as nested elements or as attributes." (But note that ExifTool will read it anyway -- ExifTool uses *very* relaxed parsing rules.)

In your last exiv2 example, combining the namespaces under a single Description is fine, but structure elements need to either be contained inside another level of Description, or the structure property itself needs to be `rdf:parseType='Resource'`. In this example, neither is true, so it isn't valid XMP.

- Phil

#35 - 21 Jan 2014 09:58 - Alan Pater

Thanks Phil, I think most of that was me messing up the dc and dwc namespaces. I've built a new version with what I think are now proper namespace separation.

It results in the following. Is this valid XMP?

```
<rdf:Description rdf:about="" xmlns:dwc="http://rs.tdwg.org/dwc/index.htm" xmlns:dc="http://purl.org/dc/elements/1.1/">
```

```
  <dwc:Taxon>
```

```
  <rdf:Description dwc:class="Vertebrata" dwc:family="Felidae" dwc:genus="Puma" dwc:kingdom="Animalia">
```

```
  <dwc:taxonRemarks>
```

```
    <rdf:Alt>
```

```
      <rdf:li xml:lang="en-US">this name ...</rdf:li>
```

```
    </rdf:Alt>
```

```
</dwc:taxonRemarks>
```

```
<dwc:vernacularName>
```

```
  <rdf:Alt>
```

```
    <rdf:li xml:lang="en-US">Cougar</rdf:li>
```

```
    <rdf:li xml:lang="es-ES">Puma</rdf:li>
```

```
  </rdf:Alt>
```

```
</dwc:vernacularName>
```

```
</rdf:Description>
```

```
</dwc:Taxon>
```

```
<dc:rights>
```

```
  <rdf:Alt>
```

```
    <rdf:li xml:lang="en-US">Alan Pater CC</rdf:li>
```

```
    <rdf:li xml:lang="es-ES">CC Alan Pater</rdf:li>
```

```
  </rdf:Alt>
```

```
</dc:rights>
```

```
</rdf:Description>
```

#36 - 21 Jan 2014 10:17 - Phil Harvey

Alan Pater wrote:

Is this valid XMP?

I don't think so. It violates section 7.9.2.4 of the XMP specification. You're still mixing nested elements (ie. taxonRemarks) with attributes (ie. class).

- Phil

#37 - 22 Jan 2014 04:19 - Phil Harvey

Phil Harvey wrote:

I don't think so.

I take this back. Maybe I have misinterpreted the specification, because this is the way that Photoshop writes dwc.

- Phil

#38 - 22 Jan 2014 04:55 - Phil Harvey

I've spent some time studying the XMP 2012 specification. I had not appreciated the significant changes made since the 2010 version. I was wrong with what I said earlier. Section 7.9.2.4 deals with a new type of structure that does not use either an inner rdf:Description or a rdf:parseType='Resource'. So this is valid, and I was wrong about what I said earlier. The restriction on not mixing nested elements and attributes applies only to this form of a structure (because, according to the spec, this form must have an empty element content).

The bottom line is that the last XMP you posted is fine.

- Phil

#39 - 22 Jan 2014 12:31 - Alan Pater

I wonder what is going on then. I've been sending my test images to Frank Bungartz at the Charles Darwin Foundation and he can't see (with the IDimager app he uses) some of the DwC I am writing with my patch of DwC. And yet it appears that IDimager and exiv2 are writing these fields in a similar manner.

exiv2.dc.dwc.i18n.jpg: Idimager can see taxonRemarks and vernacularName, but none of the other field values.

<dwc:Taxon>

```
<rdf:Description dwc:acceptedNameUsage="Tamias minimus" dwc:acceptedNameUsageID="8fa58e08-08de-4ac1-b69c-1235340b7001" dwc:class="Vertebrata" dwc:family="Felidae" dwc:genus="Puma" dwc:higherClassification="Animalia;Chordata;Vertebrata;Mammalia;Theria;Eutheria" dwc:infraspecificEpithet="concolor" dwc:kingdom="Animalia" dwc:nameAccordingTo="McCranie, J. comments" dwc:nameAccordingToID="doi:10.1016/S0269-915X(97)80026-2" dwc:namePublishedIn="Pearson O." dwc:namePublishedInID="http://hdl.handle.net/10199/7" dwc:namePublishedInYear="2059" dwc:nomenclaturalCode="ICBN" dwc:nomenclaturalStatus="nom. ambig." dwc:order="Mammalia" dwc:originalNameUsage="Gasterosteus saltatrix" dwc:parentNameUsage="Rubiaceae" dwc:parentNameUsageID="8fa58e08-08de-4ac1-b69c-1235340b7001" dwc:phylum="Chordata" dwc:scientificName="Ctenomys sociabilis" dwc:scientificNameAuthorship="(Torr.) J.T." dwc:scientificNameID="urn:lsid:ipni.org:names:37829-1:1.3" dwc:specificEpithet="concolor" dwc:subgenus="Puma" dwc:taxonConceptID="8fa58e08-08de-4ac1-b69c-1235340b7001" dwc:taxonID="8fa58e08-08de-4ac1-b69c-1235340b7001" dwc:taxonRank="subspecies" dwc:taxonomicStatus="invalid" dwc:verbatimTaxonRank="Agamospecies">
```

```
<dwc:taxonRemarks>
  <rdf:Alt>
    <rdf:li xml:lang="en-US">this name ...</rdf:li>
  </rdf:Alt>
</dwc:taxonRemarks>
```

```
<dwc:vernacularName>
  <rdf:Alt>
    <rdf:li xml:lang="en-US">Cougar</rdf:li>
    <rdf:li xml:lang="es-ES">Puma</rdf:li>
  </rdf:Alt>
</dwc:vernacularName>
```

```
</rdf:Description>
```

</dwc:Taxon>

Test_IdimagerTaxonXMP.jpg:

<dwc:Taxon>

```
<rdf:Description dwc:taxonID="test" dwc:scientificNameID="test" dwc:acceptedNameUsageID="test" dwc:parentNameUsageID="test" dwc:nameAccordingToID="test" dwc:namePublishedInID="test" dwc:taxonConceptID="test" dwc:scientificName="test" dwc:acceptedNameUsage="test" dwc:parentNameUsage="test" dwc:originalNameUsage="test" dwc:nameAccordingTo="test" dwc:namePublishedIn="test" dwc:higherClassification="test" dwc:kingdom="test" dwc:phylum="test" dwc:class="test" dwc:order="test" dwc:family="test" dwc:genus="test" dwc:subgenus="test" dwc:specific
```

```
cEpithet="test" dwc:taxonRank="test" dwc:verbatimTaxonRank="test" dwc:infraspecificEpithet="test" dwc:scientificNameAuthorship="test" dwc:nomenclaturalCode="test" dwc:taxonomicStatus="test" dwc:nomenclaturalStatus="test" dwc:taxonRemarks="test">
```

```
<dwc:vernacularName>  
  <rdf:Alt>  
    <rdf:li xml:lang="x-default">test</rdf:li>  
    <rdf:li xml:lang="en-US"/>  
    <rdf:li xml:lang="es-ES"/>  
    <rdf:li xml:lang="fr-FR"/>  
  </rdf:Alt>  
</dwc:vernacularName>
```

```
</rdf:Description>
```

```
</dwc:Taxon>
```

#40 - 23 Jan 2014 05:13 - Phil Harvey

I think this is a question for Idimager.

Photoshop reads this OK, which indicates that the XMP is well-structured.

The tags written all correspond to known tags in ExifTool (with the exception of Taxon namePublishedInYear, which I need to add).

- Phil

#41 - 23 Jan 2014 19:45 - Andreas Huggel

Alan, could you share the Exiv2 commands (or command files) you used to create your examples?

Andreas

#42 - 24 Jan 2014 11:50 - Alan Pater

- File *dwc.patch* added

- File *exiv2.dc.dwc.i18n.txt* added

- File *exiv2.dc.dwc.i18n.jpg* added

Certainly.

exiv2.dc.dwc.i18n.jpg was created using the `convert` command from `imagemagick`, and then the metadata was added from a text file using the command:

```
exiv2 -m exiv2.dc.dwc.i18n.txt exiv2.dc.dwc.i18n.jpg
```

The text file contains a list of commands such as

```
set Xmp.dwc.Record/dwc:basisOfRecord "FossilSpecimen"  
set Xmp.dwc.Taxon/dwc:vernacularName "lang=es-es Puma"
```

Also attached is the patch version used.

#43 - 20 Apr 2014 20:23 - Alan Pater

I'm still trying to figure out why `exiv2` can print pretty other nested tags but not my DwC ones. For example:

```
$ exiv2 -PXkyctl DwC-SampleImage.jpg | grep City
```

```
Xmp.iptc.CreatorContactInfo/Iptc4xmpCore:City Contact Info-City XmpText 4 Xela
```

```
$ exiv2 -PXkyctl DwC-SampleImage.jpg | grep eventID
```

```
Xmp.dwc.Event/dwc:eventID Event/dwc:eventID XmpText 4 1234
```

Instead of `Event/dwc:eventID` it should read: `Event ID`

Why doesn't it?

#44 - 25 Jun 2014 18:09 - Robin Mills

Alan

I can't answer your question about the differences.

However I've looked at your patch and test files. I believe the code in xml.cpp and properties.cpp has been submitted. I've added an additional test to bugfixes-test.sh using the your files in [r3266](#)

```
num=937a
filename=exiv2.dc.dwc.i18n.jpg
dataname=exiv2.dc.dwc.i18n.txt
diffname=exiv2.dc.dwc.i18n.diff
printf "$num " >&3
echo '----->' Bug $num '<-----' >&2
copyTestFile $filename
copyTestFile $dataname
copyTestFile $diffname
runTest exiv2 -pa $filename | sort > $num-before.txt
exiv2 -m $dataname $filename
runTest exiv2 -pa $filename | sort > $num-after.txt
diff $num-before.txt $num-after.txt > $num.txt
diff $num.txt $diffname
```

Effectively, this does:

- 1) exiv2 -pa exiv2.dc.dwc.i18n.jpg > before.txt
- 2) exiv2 -m exiv2.dc.dwc.i18n.txt exiv2.dc.dwc.i18n.jpg # apply the data file
- 3) exiv2 -pa exiv2.dc.dwc.i18n.jpg | sort > after.txt
- 4) diff before.txt after.txt

The changes feel trivial to me.

```
lcl
< Xmp.dc.language XmpBag 1 latin
---
> Xmp.dc.language XmpBag 2 latin, latin
```

Is it possible to provide a version of exiv2.dc.dwc.i18n.jpg which has no DwC data and then we can be very confident that lots of DwC data has been correctly added.

If you want me to make changes to xml.cpp and properties.cpp, could you prepare a new patch file against the current head of trunk and I will submit your code.

Robin

#45 - 25 Jun 2014 21:36 - Alan Pater

- File dwc.2014.06.25.patch added

- File blank.test.jpg added

Robin, yes, please submit this new patch. The one from back on December had a few errors, this one uses the compatible namespace. Also attached is a blank test image without any metadata.

#46 - 26 Jun 2014 02:42 - Robin Mills

- Status changed from Assigned to Resolved

Thanks, Alan. I've submitted the patch code to xml.cpp and properties.cpp. [r3267](#).

I've also updated exiv2.dc.dwc.i18n.jpg and added exiv2.dc.dwc.i18n.diff (which I forgot to submit into [r3266](#)).

I think we're complete on this, so I'll set the status to "Resolved". We'll close this issue during our review process prior to shipping Exiv2 v0.25. If anything else comes to light before we close, we can track in this issue report. Once we have closed, we'll require a new issue to track DwC.

#47 - 28 Jan 2015 19:06 - Alan Pater

As 0.25 gets closer, can I be greedy and ask for my name under Assignee on this issue?

From a comment on an other issue, I understand that will get my name in the release notes/get my name in lights ...

#48 - 28 Jan 2015 23:38 - Robin Mills

- Assignee changed from Robin Mills to Alan Pater

No problem. And you've been promoted to "Contributor". With promotion comes great responsibility. I think you can assign status and priority to issues. Use your powers wisely and may the force be with you.

Thanks for helping to develop Exiv2.

We hope that Exiv2 v0.25 will be released towards the end of February. Of course, we could be delayed by unexpected issues or team members having higher priorities in their life. Here's the current status: <http://dev.exiv2.org/boards/3/topics/1765>

Robin

#49 - 08 May 2015 16:22 - Alan Pater

- % Done changed from 30 to 100

#50 - 21 Jun 2015 16:42 - Andreas Huggel

- Status changed from Resolved to Closed

Files

DwCproperties.cpp	41.4 KB	08 Dec 2013	Alan Pater
DwC-SampleImage.jpg	52.3 KB	08 Dec 2013	Alan Pater
DwC2.txt	20.3 KB	09 Dec 2013	Robin Mills
DwC.patch	42.8 KB	09 Dec 2013	Robin Mills
dwc.patch	51 KB	21 Dec 2013	Alan Pater
xmp.cpp	29.5 KB	21 Dec 2013	Alan Pater
properties.cpp	258 KB	21 Dec 2013	Alan Pater
dwc.patch	67.9 KB	26 Dec 2013	Alan Pater
dwc.patch	72.9 KB	24 Jan 2014	Alan Pater
exiv2.dc.dwc.i18n.txt	10.1 KB	24 Jan 2014	Alan Pater
exiv2.dc.dwc.i18n.jpg	24.7 KB	24 Jan 2014	Alan Pater
dwc.2014.06.25.patch	129 KB	25 Jun 2014	Alan Pater
blank.test.jpg	13.8 KB	25 Jun 2014	Alan Pater