How do you solve a problem like multimedia?

Sharon Grant, Kate Webbink, and Marc Lambruschi

The Columbian Museum of Chicago

When the World's Columbian Exposition closed in the autumn of 1893, and its 25 million visitors had gone home, all that remained on the fairgrounds was the Palace of Fine Arts. A year later, the Columbian Museum of Chicago opened.The *Chicago Times* reported "It was, all like a memory of the fair."
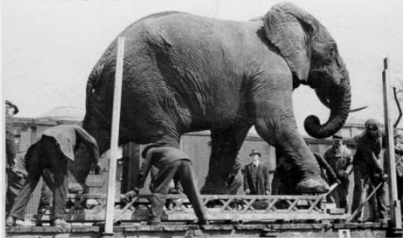(http://www.encyclopedia.chicagohistory.org/pages/450.html)

W.B. CONKEY CO CHI.

The Field Columbian Museum.

The museum's collections quickly outgrew the space in the old Palace, and in first decades of the twentieth century plans for a new museum building began to take shape. The result, a massive white marble building rising from the mud. The Field Columbian Museum opened to 8,000 visitors in May 1921.
(http://www.encyclopedia.chicagohistory.org/pages/450.html)

© 2012 – Sharon Grant, Kate Webbink, Marc Lambruschi, Mike Yoshida – The Field Museum of Natural History. Chicago. IL

Chicago Natural History Museum.

The Depression, The Great War, the attack on Pearl Harbour and another name, none of these change The Museum's goals "Accumulation and dissemination of knowledge, and the preservation and exhibition of objects illustrating art, archaeology, science and history."

TWDG 2014: Sharon Grant, Kate Webbink, Marc Lambruschi, Mike Yoshida – The Field Museum of Natural History, Chicago, IL

The Field Museum of Natural History.

Into the 21st century, a new focus on scientific research increases The Field Museum of Natural History's collections to 24 million and 186,000 square feet of storage space is created 2 floors beneath the ground.

© The Field Museum, GN91340_09d

TWDG 2014: Sharon Grant, Kate Webbink, Marc Lambruschi, Mike Yoshida – The Field Museum of Natural History. Chicago. IL

What's all this fuss about digitization?

The Impossible Dream?

Had you noticed that there is an effort underway to digitize collections?.

To capture specimen data fast enough to digitise entire collections while maintaining data quality.

To store and retrieve the assets created.

To share the information held in these collections.

How do you do this?

Can we do this?

Should we do this?

CAN WE REALIZE THE IMPOSSIBLE DREAM?

TWDG 2014: Sharon Grant, Kate Webbink, Marc Lambruschi, Mike Yoshida – The Field Museum of Natural History. Chicago. IL

Intermission

Imaging Frenzy

What to do when when …
the bunnies are multiplying and eating all the grass?

$y = 31.843e^{0.0663x}$
$R^2 = 0.5107$

TWDG 2014: Sharon Grant, Kate Webbink, Marc Lambruschi, Mike Yoshida – The Field Museum of Natural History. Chicago. IL

Storage Crisis

And we asked for more space and the answer was NO.
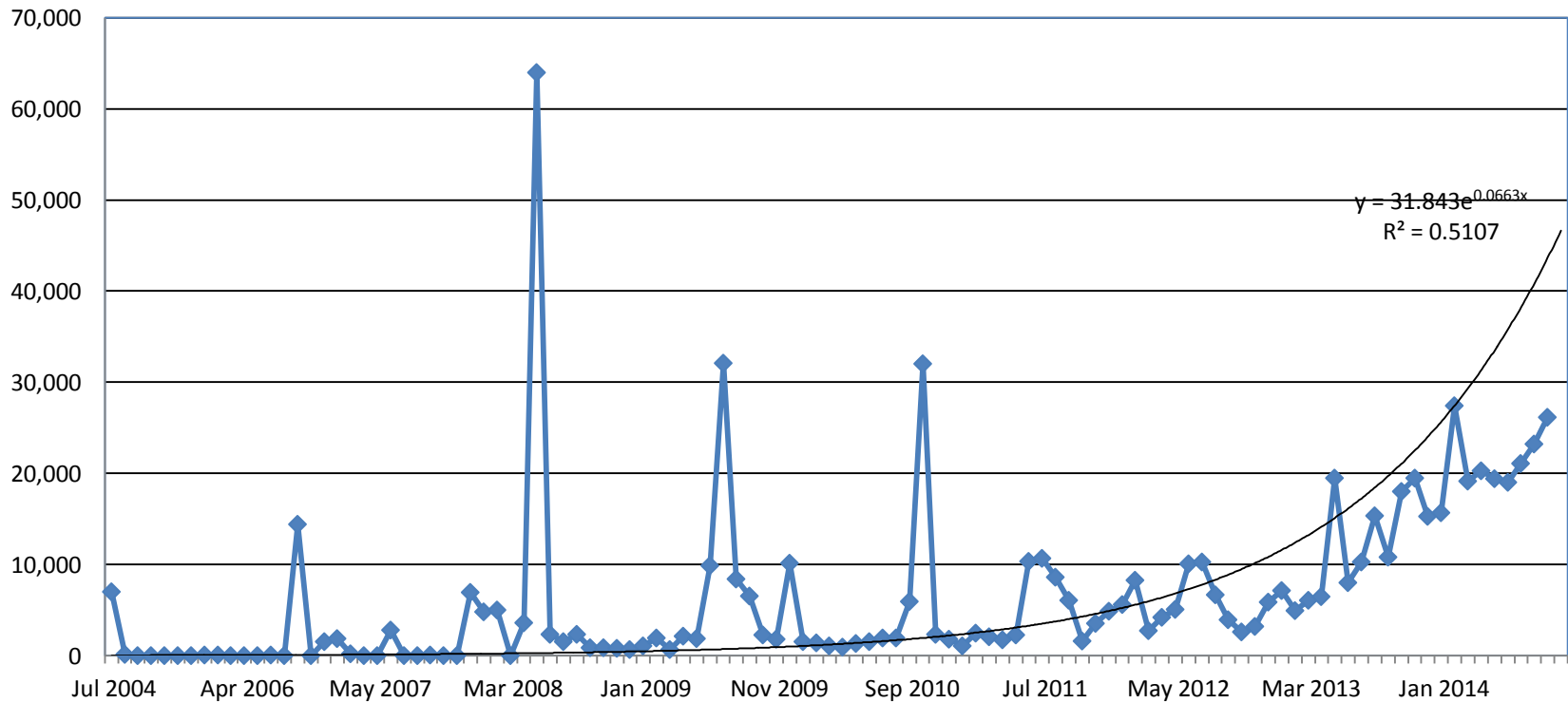

$= 0.0132e^{0.0666x}$
$R^2 = 0.4718$

Jul 2004   Fe   Apr 2014

TWDG 2014: Sharon Grant, Kate Webbink, Marc Lambruschi, Mike Yoshida – The Field Museum of Natural History. Chicago. IL

Plus our… re-indexing hell

Batch - Starting Fri May 16 22:00:01 2014

eaccessionlots... Records updated : 0
ebibliography... Records updated : 0
ecatalogue... Records updated : 10478
ecollectionevents... Records updated : 5950
econdition... Records updated : 0
econservation... Records updated : 228
edocuments... Records updated : 0
eevents... Records updated : 2
eexports... Records updated : 0
efieldhelp... Records updated : 0
efmnhrepatriation... Records updated : 20
efmnhtransactions... Records updated : 36
egazetteer... Records updated : 5
egroups... Records updated : 5
einsurance... Records updated : 0
einternal... Records updated : 136
eloans... Records updated : 0
elocations... Records updated : 0
eluts... Records updated : 60
emovements... Records updated : 0
emultimedia... Records updated : 3882
enhmstratigraphy... Records updated : 0
eparties... Records updated : 28
registry... Records updated : 118
erights... Records updated : 0
eschedule... Records updated : 0
esites... Records updated : 13
estatistics... Records updated : 0
etaxonomy... Records updated : 95
etemplate... Records updated : 0
ethesaurus... Records updated : 0
etransactions... Records updated : 0
evaluations... Records updated : 0
ewebgroups... Records updated : 0
ewebusers... Records updated : 0

Finished Sat May 17 00:28:04 2014

Reindex - Starting Sun Jun 15 01:00:01 2014

eregistry... # of records updated = 12885
eaccessionlots... # of records updated = 4402
ebibliography... # of records updated = 33081
ecatalogue... Increased size to 2883905 records # of records updated = 2873905
ecollectionevents... Increased size to 1400746 records # of records updated = 1390746
econdition... # of records updated = 7935
econservation... # of records updated = 38424
edocuments... # of records updated = 19748
eevents... Increased size to 2477 records # of records updated = 2252
eexports... Increased size to 50 records # of records updated = 5
efieldhelp... # of records updated = 5909
efmnhrepatriation... # of records updated = 10
efmnhtransactions... # of records updated = 48125
egazetteer... # of records updated = 7624
egroups... # of records updated = 3
einsurance... Increased size to 50 records # of records updated = 2
einternal... Increased size to 112222 records # of records updated = 102177
eloans... # of records updated = 18
elocations... # of records updated = 76723
eluts... Increased size to 1510404 records # of records updated = 1500404
emovements... Increased size to 50 records # of records updated = 0
emultimedia... Increased size to 638947 records # of records updated = 628947
enarratives... Increased size to 250 records # of records updated = 189
enhmstratigraphy... # of records updated = 2587
eparties... Increased size to 190197 records # of records updated = 180197
erights... Increased size to 50 records # of records updated = 11
eschedule... Increased size to 50 records # of records updated = 2
esites... Increased size to 452063 records # of records updated = 442063
estatistics... Skipped compact... # of records updated = 78612
etaxonomy... Increased size to 549590 records # of records updated = 539590
etemplate... Increased size to 50 records # of records updated = 5
ethesaurus... Increased size to 907026 records # of records updated = 897026
etransactions... Increased size to 50 records # of records updated = 0
evaluations... # of records updated = 1361
ewebgroups... Increased size to 100 records # of records updated = 84
ewebusers... Increased size to 100 records # of records updated = 82
Starting background loads..

Finished Sun Jun 15 14:26:59 2014

Reindex audit - Starting Sun Jun 15 01:00:01 2014

Starting system maintenance at Sat Sep 28 22:00:02 CDT 2013
Starting compact maintenance at Sat Sep 28 22:00:02 CDT 2013
audit...
    Checking for index updates...
    Closing database...
    Skipped compact...
    Reconfiguring database...
    Increased size to 1118773 records
    Reindexing...

Error: reindex failed to start!

    Opening database...
    Starting background loads...
Finished compact maintenance at Sat Sep 28 22:00:48 CDT 2013
Finished system maintenance at Sat Sep 28 22:01:11 CDT 2013

Finished Sun Jun 15 14:26:59 2014

Run away from the multimedia.
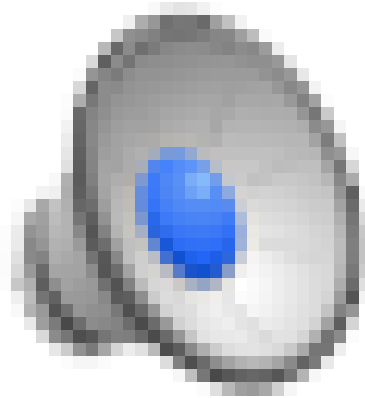
The role of multimedia in this clip is played by a fuzzy white bunny.

It wasn't quite the wild, wild west…

Intermission

HOW DO YOU SOLVE A PROBLEM LIKE MULTIMEDIA

Is a Puzzlement

File formats

Volumize with

LAYERS

Layers allow flexible editing.

Merging of images.

Addition of reference details and scale bars.

Watermarking and security.

File Formats – Layers

Each new layer added to an image increases the size of the file.

This picture consists of a blue background and on top of that a layer of conifers cut using a layer-mask in the shape of a seagull.

The blue sky used as background layer.

The greenery used as top layer.

The layer mask that was applied to the top layer. The white shape is a mask over the black.
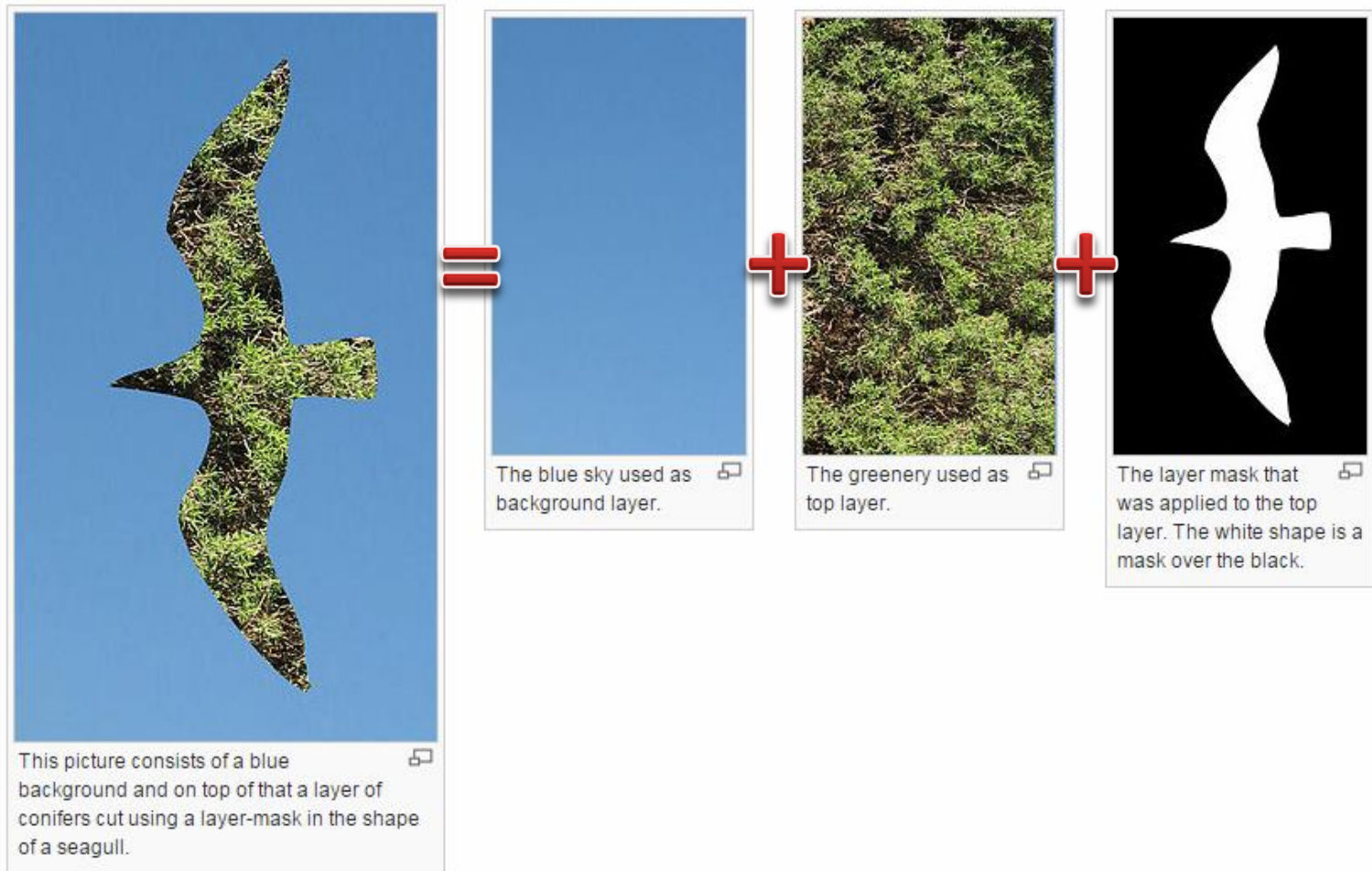
The Benefits of

BITS

More bits = More colours

More Colours = More Information

More Information = More Data

More Data = More File Space

16-bit vs 8-bit

8-bit

- Approximately 16 million tone variations

16-bit

- Approximately 281 **trillion** tone variations

Adelotopus parumpunctatus

5.0 mm

5.0 mm

• 8 Bit Color

• 16 Bit Color

Archiving S&E Imagery

	8-bit	16-bit
Color (bit-depth)		
- Layers	x MB	4x MB
+ Layers	2x MB	8x MB

x 2

File Formats – Colour Depth

8- bit approx. 16 **million** tone variations; 16-bit approx. 281 **trillion** tone variations
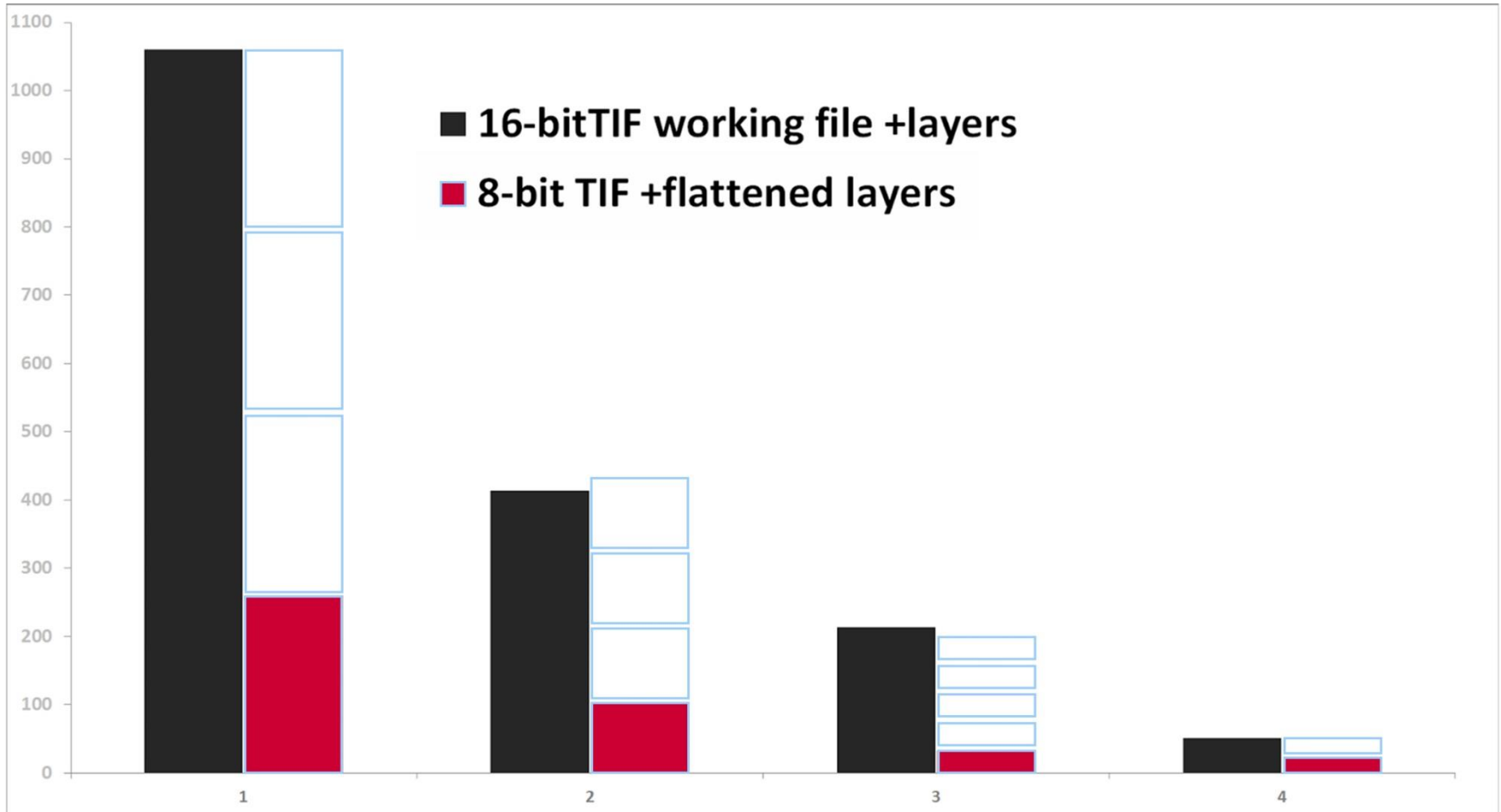
File Formats – TIFFs Vs DNGs

We archive TIFFS but they are HUGE is there anything else we could do?

File Formats – TIFFs Vs DNGs

Four Things About

DNGs

Firstly the acronym. DNG = Digital Negative (A two word TLA).

Secondly DNG is "open source" and was created by Adobe as a way to standardize RAW file formats.

Thirdly DNGs are small(er) than the TIFF that was used to create them.

Fourthly DNGs retain all the information in the TIFF AND the image is identical.

File Formats – The upshot
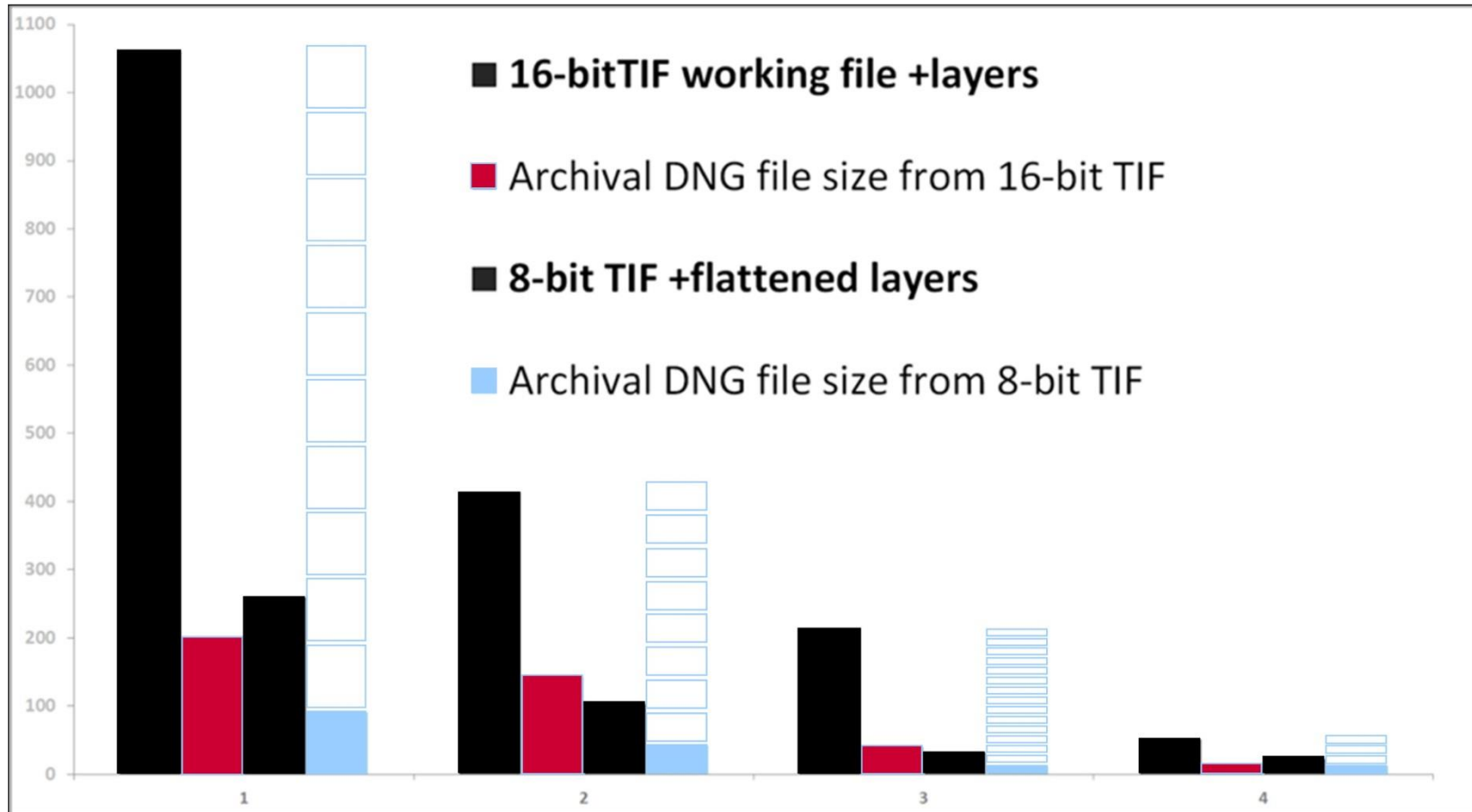
File Formats – Standards

Establishing

STANDARDS

TIFF working files; DNG archive files

DPI - 300dpi

ICC Color Profiles – AdobeRGB (& sRGB for Web)

Dimensions – < 10,000 x 10,000

Compression – None

8-bit (Approx. 16 million tone variations)

Layers – None

Storage - remember to include Backups

Ingredients:

3 copies of any important file (a primary and two backups)

on **2** different media types (such as hard drive and optical media), to protect against different types of hazards.

1 copy offsite (or at least offline).

Storage - Costs

Solution	Cost per TB for 3-2-1	Total
SAN + ExoGrid backup solution	$7,500	$2,490,000
External Hard-drives	$300	$99,600
Secret Online Sauce (shhh…)	$3,000	$996,000
Amazon Cloud Services	To hard to calculate!	Who knows?

Currently storing approx. 82 TB

150 TB is committed for digitization projects over 12-18 months (thanks NSF).

100 TB identified as under desks.

Storage - Considerations

Reimagining Storage

THE 5VS

Velocity – Speed of upload and download

Volume – Scalability and efficiency

Value – Affordability and costs

Visibility – Availability to users

Vigilance – Security and backup

Storage - Considerations

Solution	Velocity	Volume	Value	Visibility	Vigilance
SAN + ExoGrid backup solution	High	M/H*	Low*	High	High
External Hard-drives	Low	High	High	Low	Low
Secret Online Sauce (shhh…)	M/L**	High	Medium	High**	High
Amazon Cloud Services	M/L**	High	Low???	High**	High

Warning: Back of a cigarette packet assessments!

THE SERVER

The Server - Upgrading

TOBY ($18,568.41)
Live server specifications - Purchased 2009
Base Unit: DELL PowerEdge R710 with Chassis for Up to Eight 2.5-Inch Hard Drives
Processor: PowerEdge R710
Memory: 12GB Memory (6x2GB), 1333MHz Dual Ranked UDIMMs for 2
Processors, 50GB Solid State Drive SATA Mainstream 2.5in HotPlug Hard Drive,
RAID 5 for H700 or PERC 6/i Controllers, SSD Hard Drives
Hard Drive: 50GB Solid State Drive SATA Mainstream 2.5in HotPlug Hard Drive.
[Hard Drive Controller: PERC 6/i SAS RAID Controller 2x4 Connectors, Internal,
PCIe256MB Cache, x8 Chassis]

ROSS ($48,513.01)
Live server specifications - Purchased 2014
Base Unit: Cisco C240
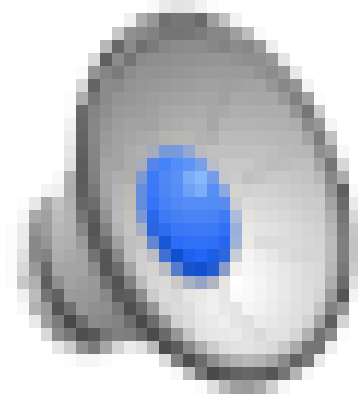Processor: Dual Intel Xeon 3.5 GHz E5-2643
Memory: 512GB of DDR3-1866-MHZ LR DIMM/PC3 RAM
Hard Drive: Dual 300GB 10K RPM SAS drives for OS and a single 785GB MLC
FusionIO drive

The Server - Upgrading

This is Rossyour oldserver

The Server – The Results

Batch inc eaudit – Start Fri Oct 3 23:00:01 CDT 2014

eaccessionlots...   Records updated   : 0
ebibliography...   Records updated   : 0
ecatalogue... Records updated   : 2213
ecollectionevents...   Records updated   : 3
econdition...   Records updated   : 12
econservation...   Records updated   : 12
edocuments...   Records updated   : 0
eevents...   Records updated   : 1
eexports...   Records updated   : 0
efieldhelp...   Records updated   : 2
efmnhrepatriation...   Records updated   :
efmnhtransactions...   Records updated   : 16
egazetteer...   Records updated   :
egroups...   Records updated   : 9
einsurance...   Records updated   : 0
einternal...   Records updated   : 704
eloans...   Records updated   : 1
elocations...   Records updated   : 0
eluts...   Records updated   : 75
emovements...   Records updated   : 0
emultimedia...   Records updated   : 1421
enarratives...   Records updated   : 0
enhmstratigraphy...   Records updated   : 0
eparties...   Records updated   : 10
eregistry...   Records updated   : 132
erights...   Records updated   : 0
eschedule...   Records updated   : 0
esites...   Records updated   : 6
estatistics...   Records updated   : 0
etaxonomy...   Records updated   : 5
etemplate...   Records updated   :
ethesaurus...   Records updated   : 0
etransactions...   Records updated   :
evaluations...   Records updated   : 0
ewebgroups...   Records updated   : 0
ewebusers...   Records updated   : 0

Finished Sat Oct 4 06:15:37 CDT 2014

Reindex – Start Sat Aug 2 22:00:01 CDT 2014

eregistry...   Number of records updated = 13321
eaccessionlots...   Number of records updated = 4405
ebibliography...   Number of records updated = 33116
ecatalogue...   Increased size to 2913319 records Number of records updated = 2903319
ecollectionevents...   Increased size to 1404107 records Number of records updated = 1394107
econdition...   Increased size to 279641 records Number of records updated = 269641
econservation...   Increased size to 44400 records Number of records updated = 40364
edocuments...   Number of records updated = 1972
eevents...   Number of records updated = 27
eexports...   Increased size to 50 records Number of records updated =
efieldhelp...   Number of records updated = 60
efmnhrepatriation...   Number of records updated = 11111
efmnhtransactions...   Number of records updated = 48531
egazetteer...   Number of records updated = 7624
egroups...   Number of records updated = 3985
einsurance...   Increased size to 50 records Number of records updated = 2
einternal...   Increased size to 1057828 records Number of records updated = 1057828
eloans...   Number of records updated = 3188
elocations...   Number of records updated = 76747
eluts...   Increased size to 1513246 records Number of records updated = 1513246
emovements...   Increased size to 50 records Number of records updated = 0
emultimedia...   Increased size to 676005 records Number of records updated = 666005
enarratives...   Increased size to 500 records Number of records updated = 259
enhmstratigraphy...   Number of records updated = 2594
eparties...   Number of records updated = 180453
erights...   Increased size to 50 records Number of records updated = 11
eschedule...   Increased size to 50 records Number of records updated = 3
esites...   Increased size to 32251 records Number of records updated = 44951
estatistics...   Number of records updated = 78
etaxonomy...   Increased size to 555903 records Number of records updated =
etemplate...   Increased size to 50 records Number of records updated = 6
ethesaurus...   Increased size to 907026 records Number of records updated = 897026
etransactions...   Increased size to 50 records Number of records updated = 0
evaluations...   Number of records updated = 1361
ewebgroups...   Increased size to 100 records Number of records updated = 84
ewebusers...   Increased size to 100 records Number of records updated = 82

Finished Sun Aug 3 01:42:06 CDT 2014

Reindex audit – Start Sun Oct 5 22:00:01 CDT 2014

eaudit...   Increased size to 48095230 records Number of records updated = 48095230
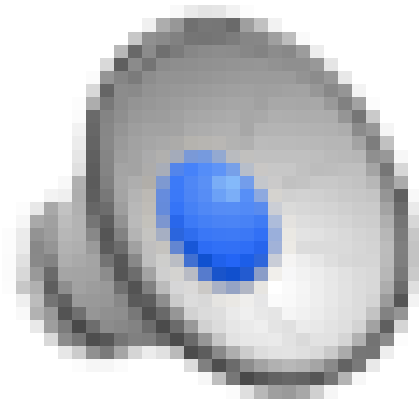
Finished Mon Oct 6 06:51:40 CDT 2014

07:15:36   03:42:05   08:51:39   13 hours per week

EMu

EMu-Magic

In the above clip please substitute EMu for the word OZ
The broom stick plays the role of digitized specimen data with problems.
The role of the faithful collection manager is played by Dorothy
How exactly do you get help some multimedia help around here?

Know thy system.

Empower your

DATABASE

Match your server to your system.

Change your workflow NOT your database.

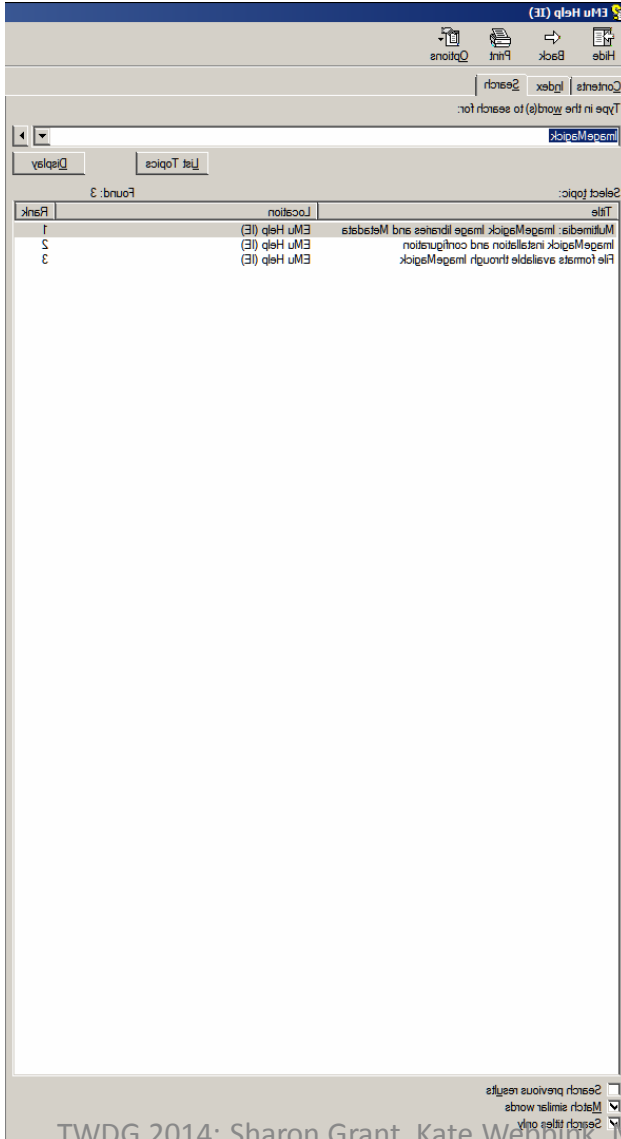Change your database NOT your workflow.

Connect to other peoples data and systems.

EMU Tools

Converting

THINGS

ImageMagick® is a software suite to create, edit, compose, or convert bitmap images. It can read and write images in a variety of **formats** (over 100) including **DPX**, **EXR**, GIF, JPEG, JPEG-2000, PDF, PNG, Postscript, SVG, and TIFF. Use ImageMagick to resize, flip, mirror, rotate, distort, shear and transform images, adjust image colors, apply various special effects, or draw text, lines, polygons, ellipses and Bézier curves. - http://www.imagemagick.org/

dcraw- (pronounced "dee-see-raw"), has become a standard tool within and without the Open Source world. It's small (about 9000 lines), portable (standard C libraries only), free (both "gratis" and "libre"), and when used skillfully, produces better quality output than the tools provided by the camera vendor. - http://www.cybercom.net/~dcoffin/dcraw/

EMu – ImageMagick

EMu - ImageMagick

Using EMu to empower your

ASSETS

EMu uses ImageMagick libraries to handle multimedia.

ImageMagick is NOTHING to do the KE Software!

ImageMagick sometimes uses other programs called delegates to handle files.
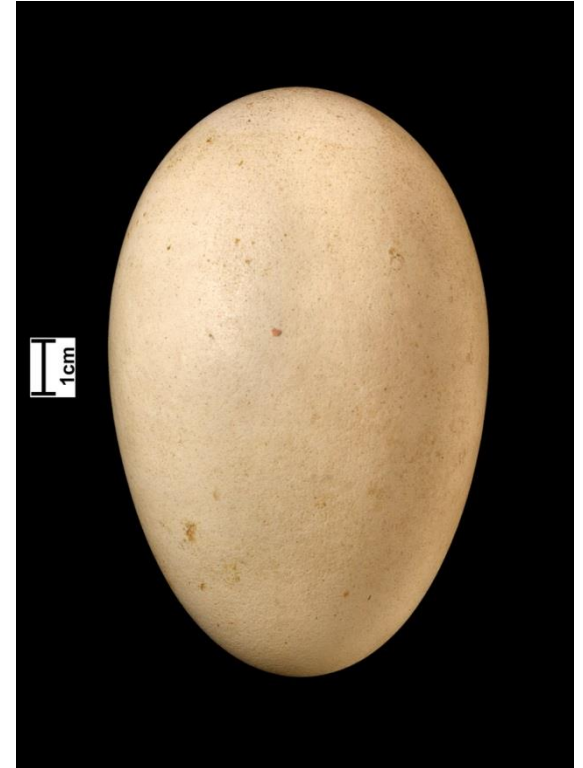
dcraw is the delegate that handles DNGs.

EMu - ImageMagick

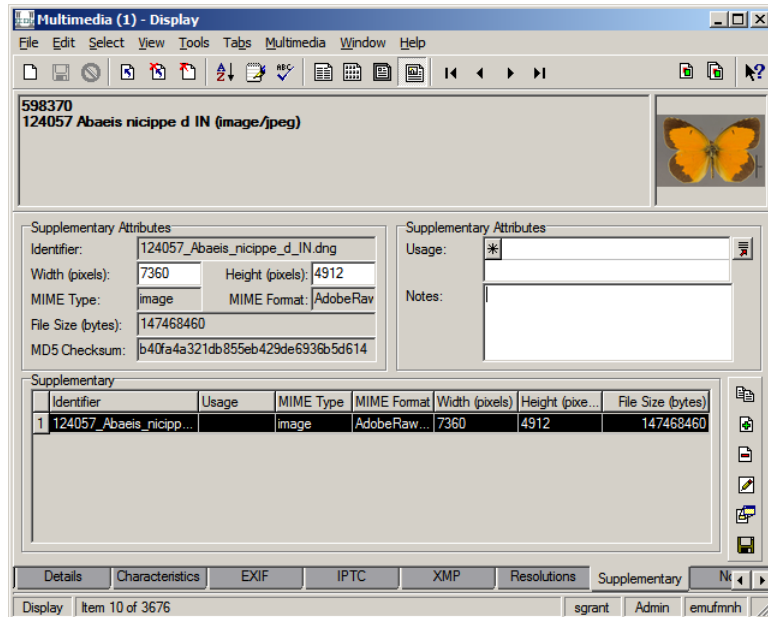

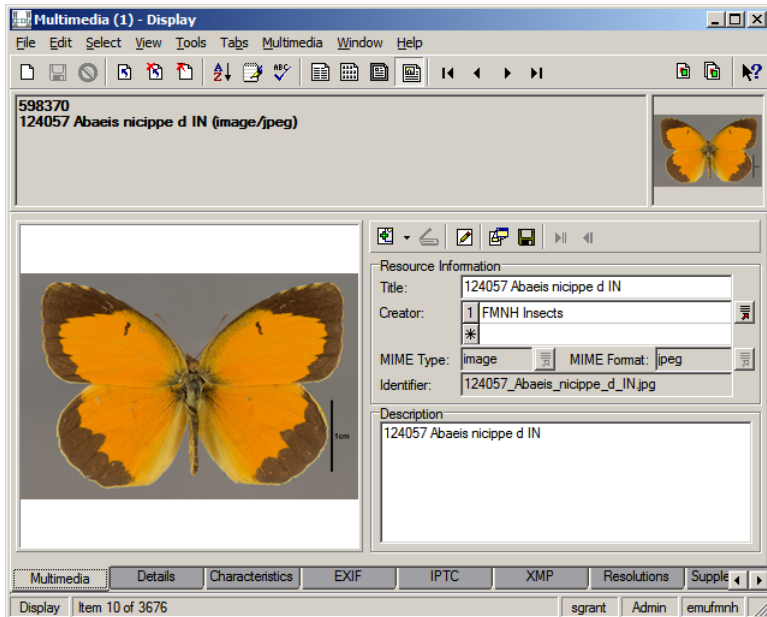TIFF->DNG   TIFF->JPG (photoshop)   DNG->JPG (EMu)

EMu - ImageMagick

7MB (Nikon raw --> DNG)	yes	10s
25MB (Canon raw --> DNG)	yes	20s
70MB (Mamiya raw --> DNG)	yes	40s
18MB raw --> DNG	yes	20s
25MB TIF --> DNG	yes	10s

3000 25MB DNGs takes 8.5 hours!

TWDG 2014: Sharon Grant, Kate Webbink, Marc Lambruschi, Mike Yoshida – The Field Museum of Natural History. Chicago. IL
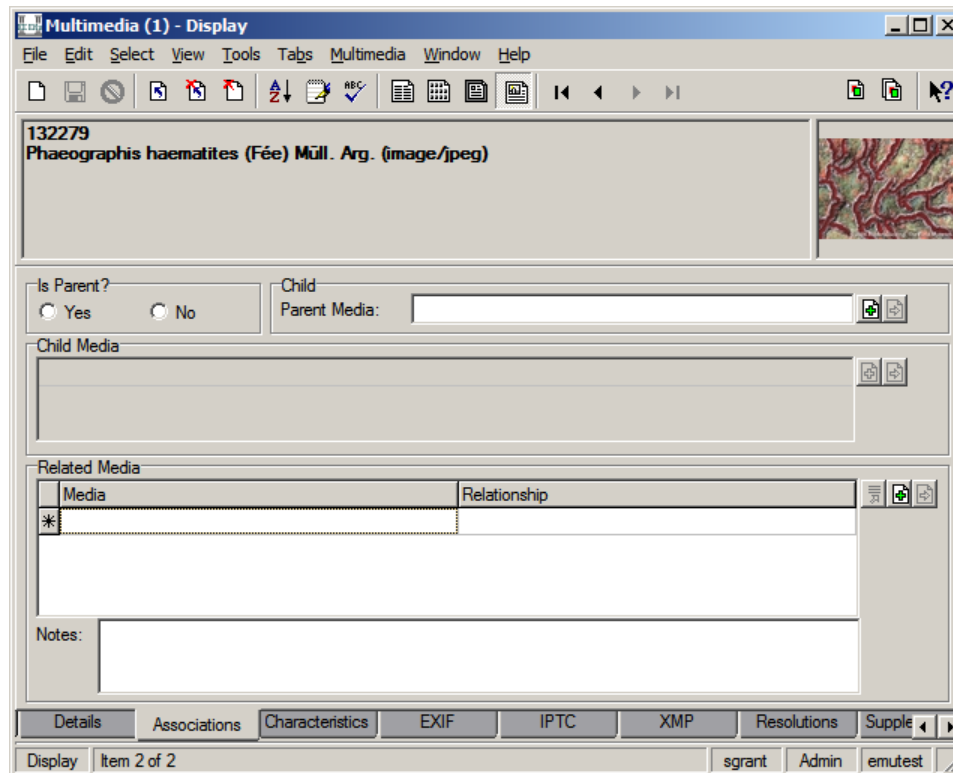
EMu - Work flow

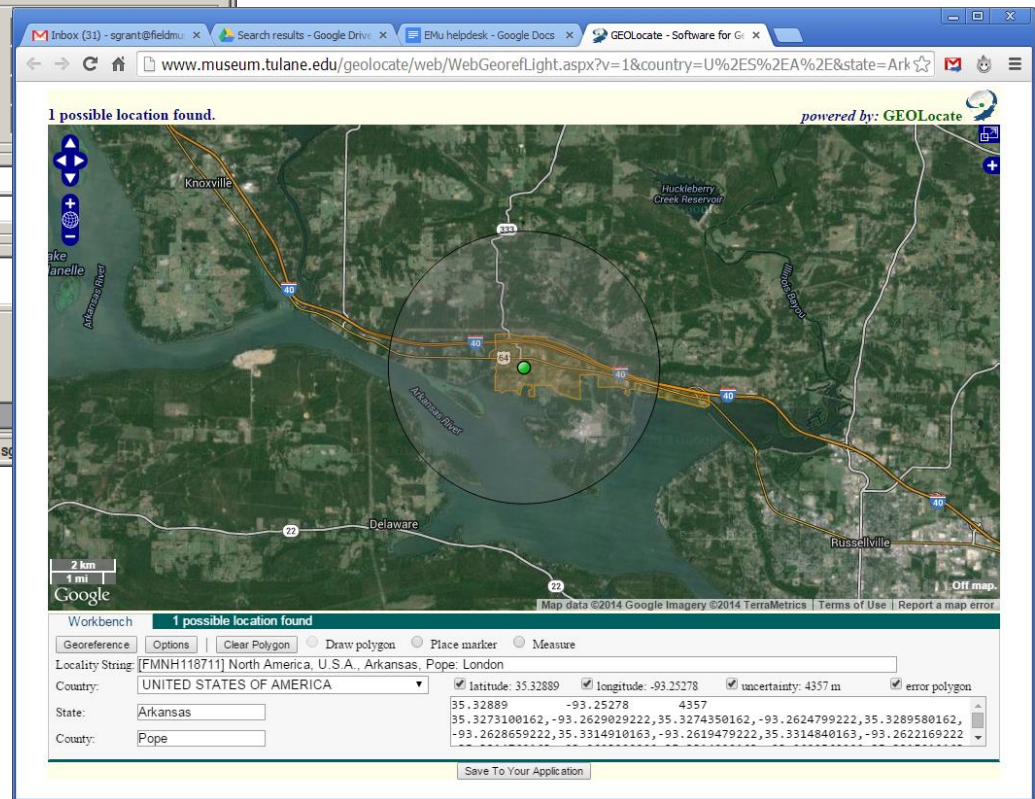The Supplementary Tab
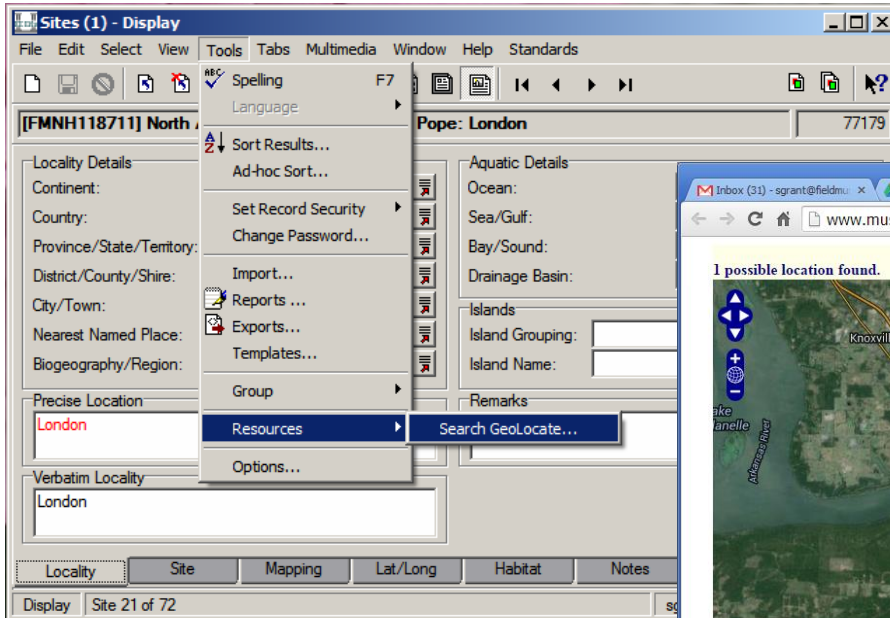
EMu - Client Changes
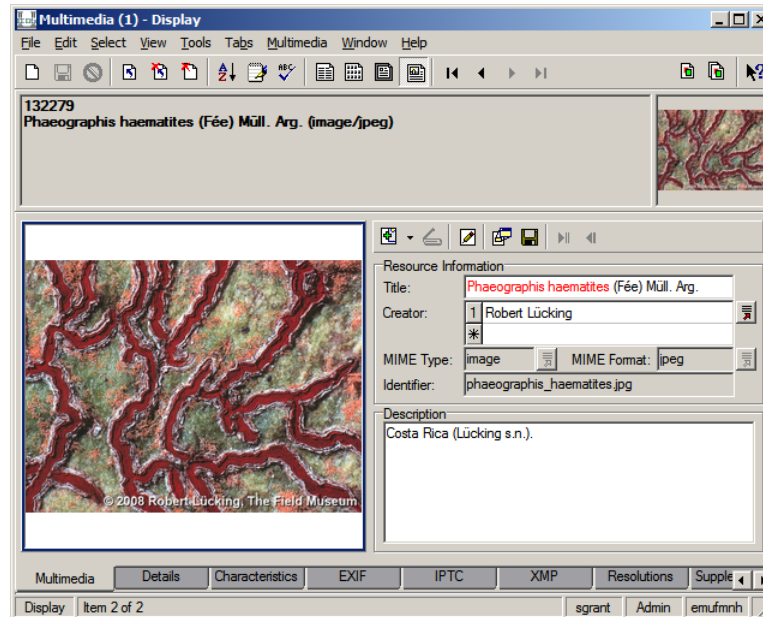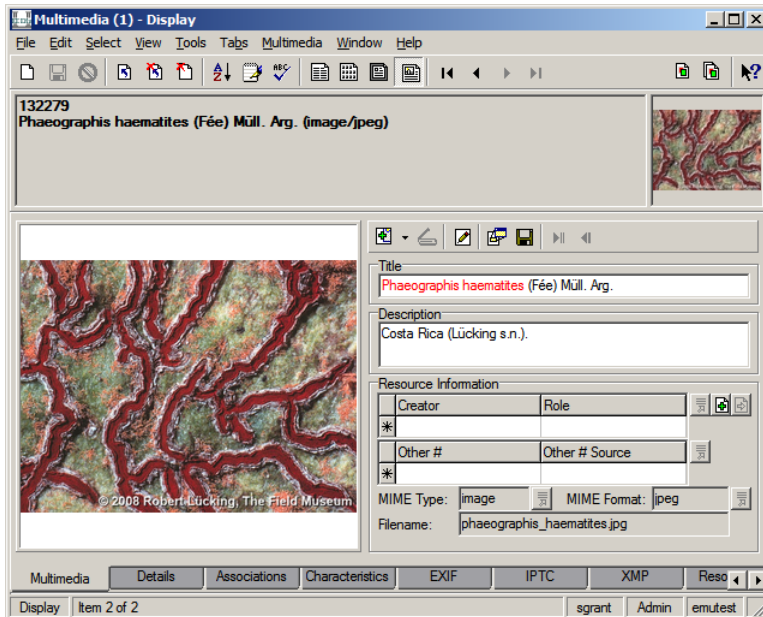
The Multimedia Tab

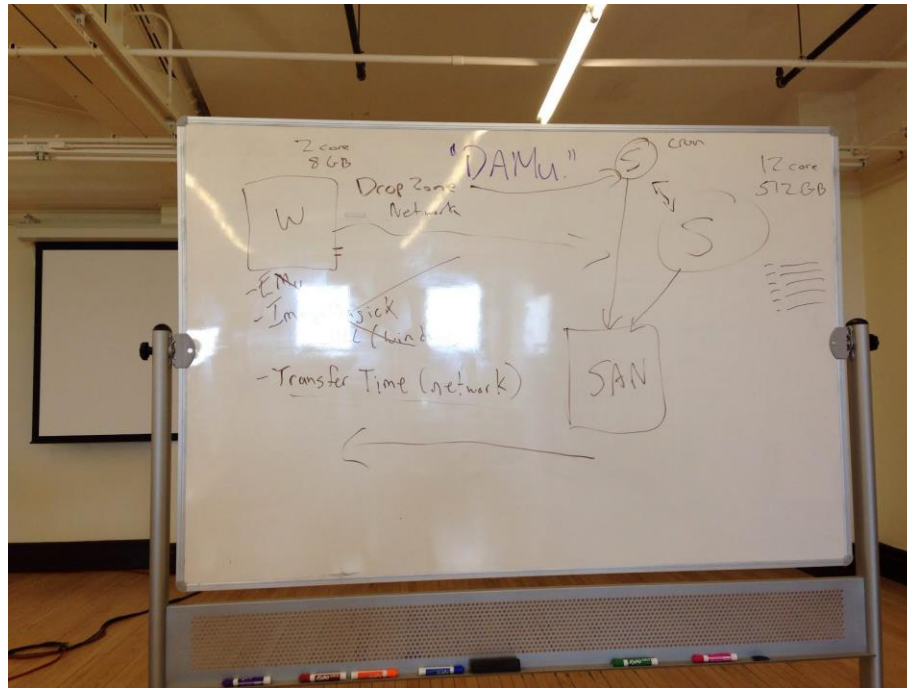# EMu - Client Changes

The Associations Tab

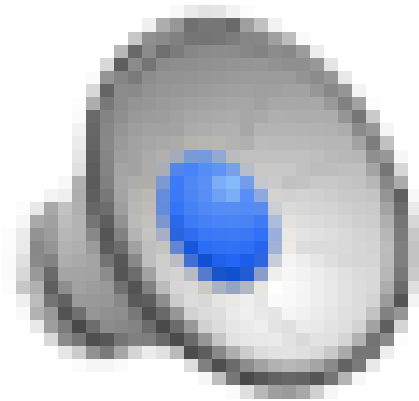# Know thy system - Emu Connections

The whiteboard challenge

Everything is easier on a whiteboard!

Finale

Thank-you

The roles of curators, collection managers, and volunteers working together to digitize specimens are played by Can-can dancers.

Thanks to Janeen Jones, Marc Lambruschi, Kate Webbink, Mike Yoshida, Rob Zschernitz and all the FMNH EMu digitizers.

Things to Read.

DIGITISING COLLECTIONS
- http://berkeleysciencereview.com/read/fall-2011/digitizing-the-drawers/
- http://zookeys.pensoft.net/articles.php?id=2916
- http://soyouthinkyoucandigitize.wordpress.com/
- https://www.facebook.com/pages/Insects-Plants-and-Parasites-Digitizing-Natural-History-Collections/330373400373054
- https://www.idigbio.org/content/gbif-training-manual-1-digitization-natural-history-collections-data
- http://images.library.amnh.org/digital/
- http://mw2013.museumsandtheweb.com/paper/building-cybercabinets-guidelines-for-online-access-to-digital-natural-history-collections/
- https://journals.ku.edu/index.php/jbi/article/view/3992
- http://www.gbif.org/resources/2775

FILE FORMATS
- http://www.phonearena.com/news/Raw-DNG-vs-JPEG-on-a-smartphone-comparison-images_id58538
- http://www.rocktheshotforum.com/2012/09/14/three-things-to-know-about-dng/
- http://www.mosaicarchive.com/2013/05/01/the-raw-truth-about-dng/
- http://photo.stackexchange.com/questions/10208/how-many-colors-and-shades-can-the-human-eye-distinguish-in-a-single-scene
- http://helpx.adobe.com/photoshop/using/layer-basics.html
- http://en.wikipedia.org/wiki/Layers_(digital_image_editing)
- http://www.phonearena.com/news/Raw-DNG-vs-JPEG-on-a-smartphone-comparison-images_id58538
- http://www.rocktheshotforum.com/2012/09/14/three-things-to-know-about-dng/
- http://www.barrypearson.co.uk/articles/dng/respectability.htm
- http://blogs.loc.gov/digitalpreservation/2014/10/the-library-of-congress-wants-you-and-your-file-format-ideas/