



Notes on Spotlight

Thursday, 17 June 2010

Depending on the file, spotlight returns differing attributes. We will use these attributes to catalogue the assets.

Points of interest

- ★ Orientation appears from time to time in the attributes. 0 = landscape 1 = portrait
- ★ The values in the “Kind” attribute is quite in-consistent. Sometimes the acronym is expanded in the case of png but not with jpeg.
- ★ The meta data in fonts is a lot more consistent

Attributes that are not preserved

- ★ Spotlight comments
- ★ Created at Modified at

Finding out classification

In order to decide on general classification of files, we will phase the kMDItemContentTypeTree

This will be a challenge for obscure files like *rack.ru* as there are only two nodes on the content type tree “public.data”, “public.item”

Attributes from a variety of file types

This list is not exhaustive. It aims to give us an idea of the types of attributes we can expect to pull back from spotlight.

Kind (from Spotlight)	Notes	Other attributes	kMDItemContentTypeTree
JPEG image	“Save as jpeg” Should be able to pull out additional info such as “date digitized”	Bits per sample (32) Pixel width/height Size Resolution width/height Colour profile Alpha channel Color space Content created	public.jpeg public.image public.data public.item public.content
	Attributes from cameras. On an iPhone 3GS you will have longitude and latitude	Focal Device make Content creator Focal length Orientation Exposure time F number White balance ISO speed Flash (0 or 1) Exposure mode Red eye	

Kind (from Spotlight)	Notes	Other attributes	kMDItemContentTypeTree
Portable Network Graphics image		Pixel count Size Resolution w/h Orientation Alpha channel 1 Colour space	public.png public.image public.data public.item public.content
Portable Document Format (PDF)	Measurements recorded in points. Which is 72 points per inch (postscript) 2.54 centimetres per inch	Content creator (software) page width (in points) Pages (count) Security Version (pdf version) Encoding software	com.adobe.pdf public.data public.item public.composite-content public.content
	With a password protected pdf, only attributes pages (count) others aren't accessible	Security Password Encrypted	
Adobe Photoshop file	Lots of information. Also includes any camera information if the image was originally taken on a digital camera	Resolution width 300 Resolution height 300 Bits per sample Colour space Width/height in pixels Alpha channel (1) Orientation 1 layers (array of quoted strings) Content creator (software name and version) Colour profile (string)	com.adobe.photoshop-image public.image public.data public.item public.content

Kind (from Spotlight)	Notes	Other attributes	kMDItemContentTypeTree
Adobe Illustrator Document	Unfortunately doesn't return dimensions or colour space		com.adobe.illustrator.ai-image public.image public.data public.item public.content
InDesign® CS4 Document	No useful meta data. Hopefully this can be changed with a 3 rd party plugin		public.data public.item
Document	Imported a CS document. mdimport did return... kMDItemKind = InDesign \U00ae CS Document" when debug was on 2		public.unix-executable public.data public.item public.executable
SVG Image	public.svg-image		public.svg-image public.image public.data public.item public.content
Encapsulated PostScript	com.adobe.encapsulated-postscript	Content creator (name of software) Page width Page height Version (of postscript)	com.adobe.encapsulated-postscript com.adobe.postscript public.data public.item public.composite-content public.content"

InDesign®

Kind (from Spotlight)	Notes	Other attributes	kMDItemContentTypeTree
OpenType font	public.opentype-font	Copyright Fonts (list of weights) Family name Style name	public.opentype-font public.font public.data public.item
Datafork TrueType font	com.apple.truetype-datafork-suitcase-font	Full name Postscript name Publishers Version Authors	com.apple.truetype-datafork-suitcase-font public.truetype-font public.font public.data public.item
Numbers Document	com.apple.iwork.numbers.sffnumbers	Fonts (array of font names) Keywords, Author, Title if entered into the info panel. (otherwise is an array of empty comma's)	com.apple.iwork.numbers.sffnumbers com.pkware.zip-archive public.data public.item com.apple.bom-archive public.archive public.zip-archive public.composite-content public.content
Pages Publication	08 pages doc		com.apple.iwork.pages.pages com.apple.package public.directory public.item public.composite-content public.content

Kind (from Spotlight)	Notes	Other attributes	kMDItemContentTypeTree
Pages Publication	09 pages doc		com.apple.iwork.pages.sffpages com.pkware.zip-archive public.data public.item com.apple.bom-archive public.archive public.zip-archive public.composite-content public.content
Keynote Document	09 keynote	Oddly no fonts used or other iWork attributes	com.apple.iwork.keynote.sffkey com.pkware.zip-archive public.data public.item com.apple.bom-archive public.archive public.zip-archive public.presentation public.composite-content public.content
Microsoft Word 97 - 2004 document	com.microsoft.word.doc		com.microsoft.word.doc public.data public.item public.composite-content public.content

Kind (from Spotlight)	Notes	Other attributes	kMDItemContentTypeTree
Microsoft Word document	docx No word count or number of pages	Authors Title	org.openxmlformats.wordprocessingml.document org.openxmlformats.openxml public.zip-archive com.pkware.zip-archive public.data public.item com.apple.bom-archive public.archive public.composite-content public.content
Microsoft Excel 97-2004 workbook			com.microsoft.excel.xls public.data public.item public.composite-content public.content
Microsoft PowerPoint 97 - 2004 presentation			com.microsoft.powerpoint.ppt public.data public.item public.presentation public.composite-content public.content
Rich Text Format			public.rtf public.text public.data public.item public.content

Kind (from Spotlight)	Notes	Other attributes	kMDItemContentTypeTree
Rich Text Format with Attachments		Number of items. Has a count of attachement	com.apple.rtf com.apple.package public.directory public.item public.composite-content public.content
ZIP archive (Actually an Open office presentation)	On the open office website they say the new version comes with a quicklook plugin		<div> org.oasis-open.opendocument.presentation org.oasis-open.opendocument public.zip-archive com.pkware.zip-archive public.data public.item com.apple.bom-archive public.archive public.composite-content public.content </div> <div> org.oasis-open.opendocument.spreadsheet org.oasis-open.opendocument public.zip-archive com.pkware.zip-archive public.data public.item com.apple.bom-archive public.archive public.composite-content public.content </div>

Kind (from Spotlight)	Notes	Other attributes	kMDItemContentTypeTree
			org.oasis-open.opendocument.text org.oasis-open.opendocument public.zip-archive com.pkware.zip-archive public.data public.item com.apple.bom-archive public.archive public.composite-content public.content
MPEG-4 movie		Media types (Sound, Video) Codecs (ACC, "H.264") Pixel width Pixel height Duration	public.mpeg-4 public.movie public.audiovisual-content public.data public.item public.content
Flash video	Basically no additional meta data short of size.		com.adobe.flash.video public.movie public.audiovisual-content public.data public.item public.content
MP3 Audio File		Album genre etc Bitrate Duration appears to be recorded in seconds	public.mp3 public.audio public.audiovisual-content public.data public.item public.content

Kind (from Spotlight)	Notes	Other attributes	kMDItemContentTypeTree
Style Sheet	Note doesn't say "Cascading Style Sheet"		com.apple.dashcode.css public.plain-text public.text public.data public.item public.content
HTML document		Title is pulled from the <title> in the head Where from...	public.html public.text public.data public.item public.content
Web archive		Content creator (Cocoa HTML writer)	com.apple.webarchive public.data public.item public.composite-content public.content
Email Message			com.apple.mail.email public.data public.item public.email-message public.message
Chat transcript	iChat transcript. Not sure what a MSN one would be like	Author address (handle) Instant message addresses (handles) Authors (names of those in chat) Delivery type. (AIM)	com.apple.ichat.transcript public.message public.data public.item

Kind (from Spotlight)	Notes	Other attributes	kMDItemContentTypeTree
JavaScript source	Says executable at base of content type tree	Where from	com.apple.dashcode.javascript com.netscape.javascript-source public.source-code public.plain-text public.text public.data public.item public.content public.executable
Shockwave Flash Movie	Minimal attributes, regardless if swf is heavily scripted or timeline animation		public.data public.item
Ruby source			public.ruby-script public.shell-script public.script public.source-code public.plain-text public.text public.data public.item public.content
vCard	If the card contains multiple people, title is pulled from the first person	Title from the name of the person	public.vcard public.data public.item public.contact