School of Computing
IT8003 Digital Forensics and Investigation

Practical 3A: Documents

Introduction

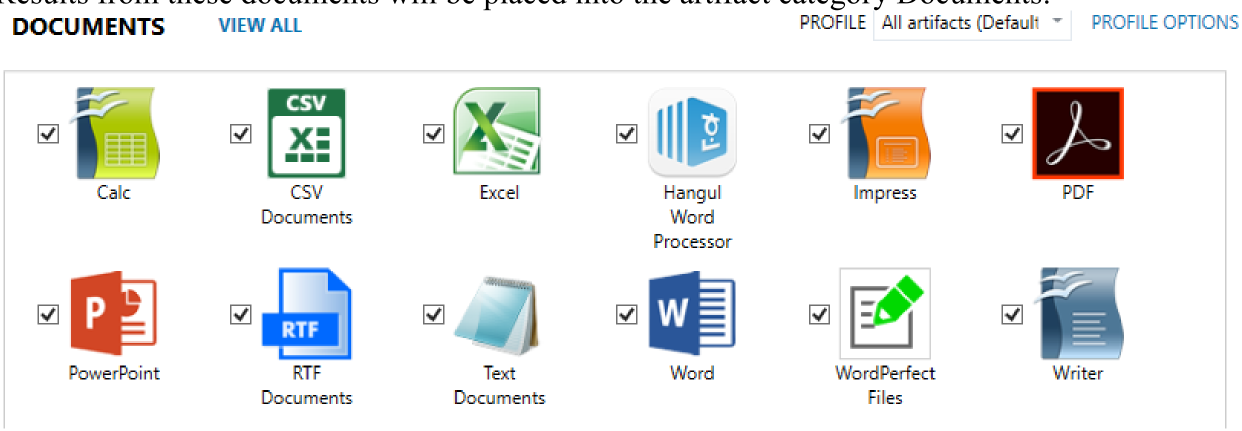
Documents are files that maybe relevant to what the subject is working on or tampering with. It can provide **metadata** information as to when the file was created or who is the author of the file. At times this is crucial especially when you are dealing with corporates company that may have contracts or sensitive documents.

Learning Objective

In this lesson, students will take part in lectures, instructor-led exercises, and student practical exercises to gain an understanding of the differing views of documents, the metadata of files and how to access AXIOM Examine's built in help and reference capabilities as well as the Artifact reference. Students will also explore the ability to search documents and metadata via the filters bar in AXIOM. At the conclusion of this lesson, students will be able to identify, discuss, and utilize Magnet AXIOM to search the data and metadata of files from AXIOM. Students will also be able to utilize the built in help and Artifact Reference and be able to utilize the filters and keywords to search document content.

Magnet Axiom Process Documents

AXIOM Process will search for and categorize a number of different types of documents. Results from these documents will be placed into the artifact category Documents:



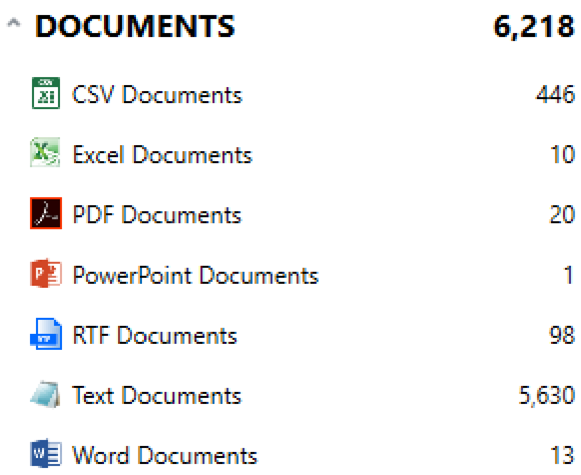
DOCUMENTS VIEW ALL PROFILE All artifacts (Default PROFILE OPTIONS

☑ Calc ☑ CSV Documents ☑ Excel ☑ Hangul Word Processor ☑ Impress ☑ PDF

☑ PowerPoint ☑ RTF Documents ☑ Text Documents ☑ Word ☑ WordPerfect Files ☑ Writer

Figure 3-3-1: Axiom Process Documents Artifacts

AXIOM Examine will provide various viewing options for Documents and their associated metadata. After processing, AXIOM Examine will present all documents in the document artifact category with a number denoting the quantity of document artifacts recovered.



DOCUMENTS	6,218
CSV Documents	446
Excel Documents	10
PDF Documents	20
PowerPoint Documents	1
RTF Documents	98
Text Documents	5,630
Word Documents	13

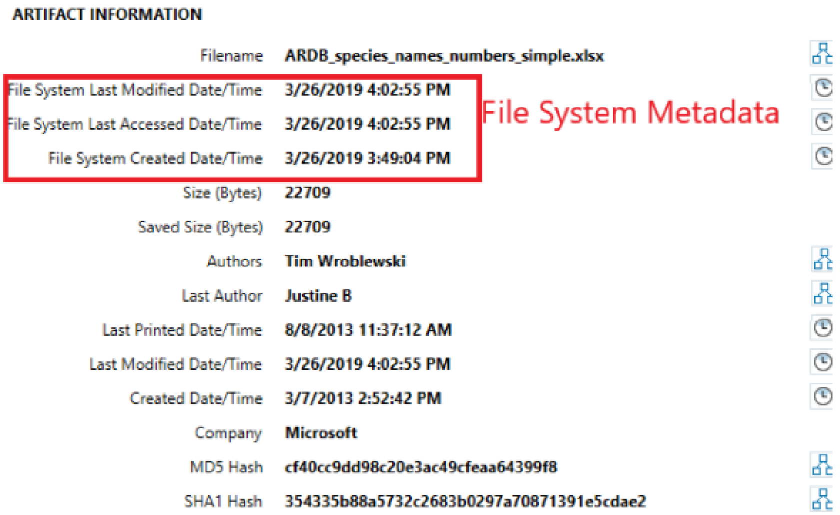
Figure 3-3-2: Documents Artifacts

The default view of artifacts in the Documents category will allow the Contents Pane to display documents much like they would appear in the original application used to create or view them. The Preview card has a built-in viewer. PowerPoint slides can be scrolled through; PDFs with multiple pages can be scrolled through; Word documents with embedded graphics may also be viewed. PDFs will often render with two preview cards: one will show the normal view of the document and the other will show a "filtered text" type of view with no graphics.

Viewing Documents

The default view of artifacts in the Documents category will show column data in the Evidence Pane (Column View) and a preview of the document contents in the Contents Pane. The following columns will be available in the Column View:

Component	Description
Filename	The name of the document.
File System Created Date/Time (UTC)	The date and time the file was created on the file system. This is also known as File System Metadata. This information is parsed from the $MFT entry for the file.
File System Last Accessed Date/Time (UTC)	The date and time the file was last accessed on the file system. This is also known as File System Metadata. This information is parsed from the $MFT entry for the file.
File System Last Modified Date/Time (UTC)	The date and time the file was last modified on the file system. This is also known as File System Metadata. This information is parsed from the $MFT entry for the file.
Size (Bytes)	The size of the file in bytes.
Saved Size (Bytes)	The size of the file in bytes that was recovered. Extremely large documents may not be fully recovered.
File	The actual file contents.

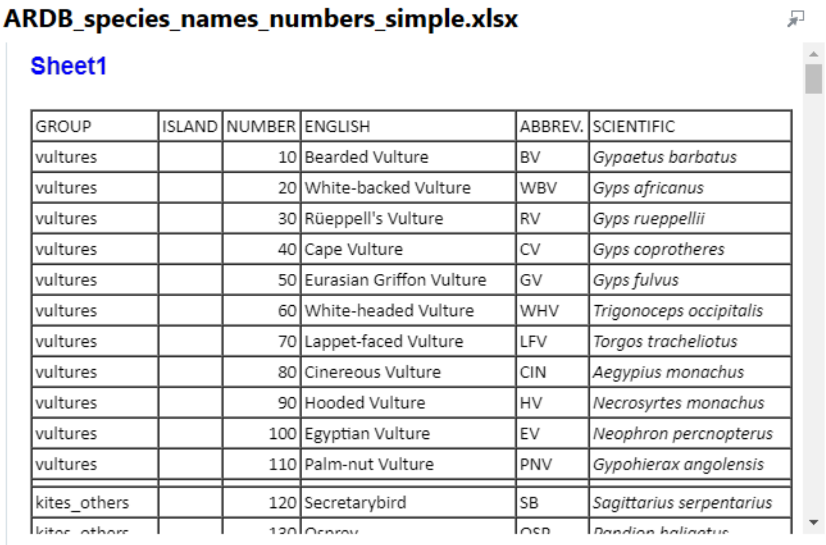


ARTIFACT INFORMATION

Filename	ARDB_species_names_numbers_simple.xlsx
File System Last Modified Date/Time	3/26/2019 4:02:55 PM
File System Last Accessed Date/Time	3/26/2019 4:02:55 PM
File System Created Date/Time	3/26/2019 3:49:04 PM
Size (Bytes)	22709
Saved Size (Bytes)	22709
Authors	Tim Wroblewski
Last Author	Justine B
Last Printed Date/Time	8/8/2013 11:37:12 AM
Last Modified Date/Time	3/26/2019 4:02:55 PM
Created Date/Time	3/7/2013 2:52:42 PM
Company	Microsoft
MD5 Hash	cf40cc9dd98c20e3ac49cfeaa64399f8
SHA1 Hash	354335b88a5732c2683b0297a70871391e5cdae2

File System Metadata

Figure 3-3-3: File System Metadata



ARDB_species_names_numbers_simple.xlsx

Sheet1

GROUP	ISLAND	NUMBER	ENGLISH	ABBREV.	SCIENTIFIC
vultures		10	Bearded Vulture	BV	*Gypaetus barbatus*
vultures		20	White-backed Vulture	WBV	*Gyps africanus*
vultures		30	Rüeppell's Vulture	RV	*Gyps rueppellii*
vultures		40	Cape Vulture	CV	*Gyps coprotheres*
vultures		50	Eurasian Griffon Vulture	GV	*Gyps fulvus*
vultures		60	White-headed Vulture	WHV	*Trigonoceps occipitalis*
vultures		70	Lappet-faced Vulture	LFV	*Torgos tracheliotus*
vultures		80	Cinereous Vulture	CIN	*Aegypius monachus*
vultures		90	Hooded Vulture	HV	*Necrosyrtes monachus*
vultures		100	Egyptian Vulture	EV	*Neophron percnopterus*
vultures		110	Palm-nut Vulture	PNV	*Gypohierax angolensis*
kites_others		120	Secretarybird	SB	*Sagittarius serpentarius*
kites_others		130	Osprey	OSP	*Pandion haliaetus*

Figure 3-3-4: Contents Preview

Connections Explorer

AXIOM Examine can identify **connections** between the **artifacts within a case**. By default, these connections are not built automatically, and the process must be started manually using the **Tools -> Build Connections** option. Alternatively, the connections can be set to automatically rebuild every time new evidence is added to the case by enabling the **Automatically build connections** option in **Tools -> Settings -> Connections**. Connection information is collected from all evidence items in the case regardless of whether it originates from mobile devices, computer-based devices, or even cloud-based evidence. AXIOM Examine then builds the connections, and links between the artifacts are identified. The Connections explorer displays these connections visually, making it faster and easier to identify and understand how various pieces of the investigative puzzle fit together. With the ever-growing mountain of evidence examiners must deal with on a day-to-day basis, Connections provides a way of connecting the dots and identifying key related information in a more expedient manner. The Connections explorer will help examiners establish the who, what, when, where, why, and how of the investigation, and it is expected that the examiner will return to the Connections explorer many times throughout the life of the case.

Component	Regular Expressions
WHO	Who was involved? Understanding who owns a suspect file; who put it in that location; who, if anyone, has looked at or executed the file (depending on the file type); who deleted it; who emailed/transferred it; who did they email/transfer it to; who was using the machine at the time the offense occurred; and who else has been using the machine are all questions that could help answer the key question – **Who was involved?**
WHAT	What happened? Understanding what other files, if any, this file is related to; what other files have the same hash regardless of filename; what applications have been used to create, modify, and send the file; what additional information does the metadata provide (Word docs - when was it last printed, Pictures – What camera was used); what other files were stored in the same folder/on the same device; and what was the sequence of events, are all questions that could help answer the key question – **What happened?**
WHEN	When did it occur? Understanding when a picture was taken (EXIF data);

	when was this file viewed, emailed/shared/transferred, when was this file deleted, when was this file executed or last accessed, are all questions that could help answer the key question – **When did it occur?**
WHERE	Where did it take place? Understanding where else a file is located, was it saved locally, to other devices, or to the cloud; where was it downloaded from; where was it distributed to; are there logs to show where a device been used, are all questions that could help answer the key question - **Where did it take place?**
WHY	Why did it happen? The content of correspondence in the form of chat, email, instant messaging communications etc.; or the content of machine activity logs could help answer the key question – **Why did it happen?**
HOW	How did it happen? How did this file get onto this device; how was the file shared with other people; how did this person communicate with other key people; the content of correspondence in the form of chat, email, instant messaging communications etc. could all help answer the key question – **How did it happen?**

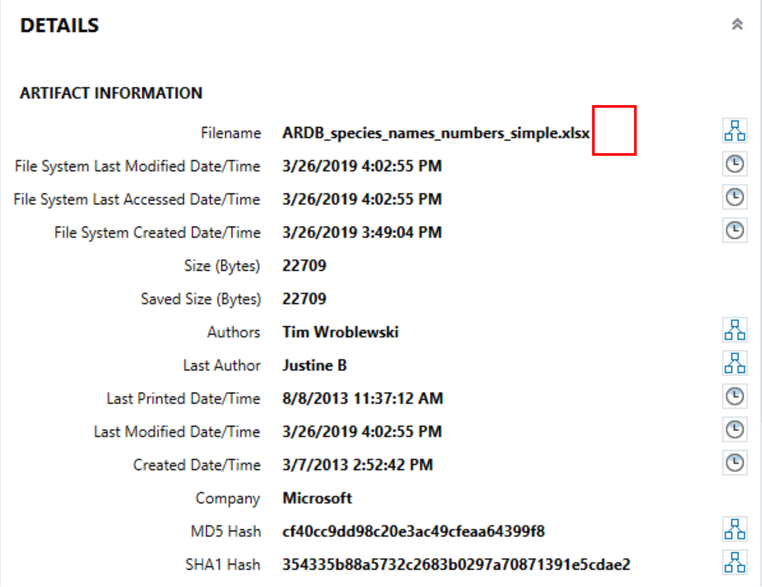
Once the connections have been built, AXIOM Examine displays a CONNECTIONS icon beside any artifact attribute that has been connected in some way. This could be the filename, hash value, metadata field etc. Clicking the CONNECTIONS icon automatically switches AXIOM Examine to the Connections explorer with the selected artifact attribute as the PRIMARY NODE.

There are four types of nodes within the Connections explorer:

1. **PRIMARY NODES** are displayed in HOT PINK. This is the anchor point from which the connections are being made. In the Artifacts or File system explorers, selecting a CONNECTIONS icon for a specific artifact attribute switches AXIOM Examine to the Connections explorer with that artifact attribute set as the primary node. Within the Connections explorer double-clicking any node sets it as the primary node.
2. **DIRECT NODES** are displayed in BLUE. These are artifact attributes with a direct connection to the primary node. To view only connections between a primary node and a direct node, click the direct node.
3. **SELECTED NODES** are displayed in TEAL. When a direct node is selected it becomes a selected node. The matching results displayed in the Connections explorer refresh to display only artifacts that contain both attributes of the primary and selected node e.g. filename and application name. When a direct node becomes a selected node, indirect connections come into focus.
4. **INDIRECT NODES** are displayed in GREY. When a direct node becomes a selected node all other direct connections to the primary node become indirect nodes and turn grey. All direct connections to the selected node are also now displayed as indirect nodes.

CONNECTORS are the lines representing connections between two nodes. Types of connections include: shares partial path, accessed with, transferred to, source, etc. Connections Explorer can only be launched from Artifact Explorer. For each item for which there is connection information available, a small connection icon will be next to the item.

Go to Excel documents and select the document
ARDB_species_names_numbers_simple.xlsx.
Click on the connections icon next to the file name:



DETAILS

ARTIFACT INFORMATION

Filename	ARDB_species_names_numbers_simple.xlsx
File System Last Modified Date/Time	3/26/2019 4:02:55 PM
File System Last Accessed Date/Time	3/26/2019 4:02:55 PM
File System Created Date/Time	3/26/2019 3:49:04 PM
Size (Bytes)	22709
Saved Size (Bytes)	22709
Authors	Tim Wroblewski
Last Author	Justine B
Last Printed Date/Time	8/8/2013 11:37:12 AM
Last Modified Date/Time	3/26/2019 4:02:55 PM
Created Date/Time	3/7/2013 2:52:42 PM
Company	Microsoft
MD5 Hash	cf40cc9dd98c20e3ac49cfeaa64399f8
SHA1 Hash	354335b88a5732c2683b0297a70871391e5cdae2

Figure 3-3-11: Connections Icon

Exercise 1. Viewing Documents – Metadata

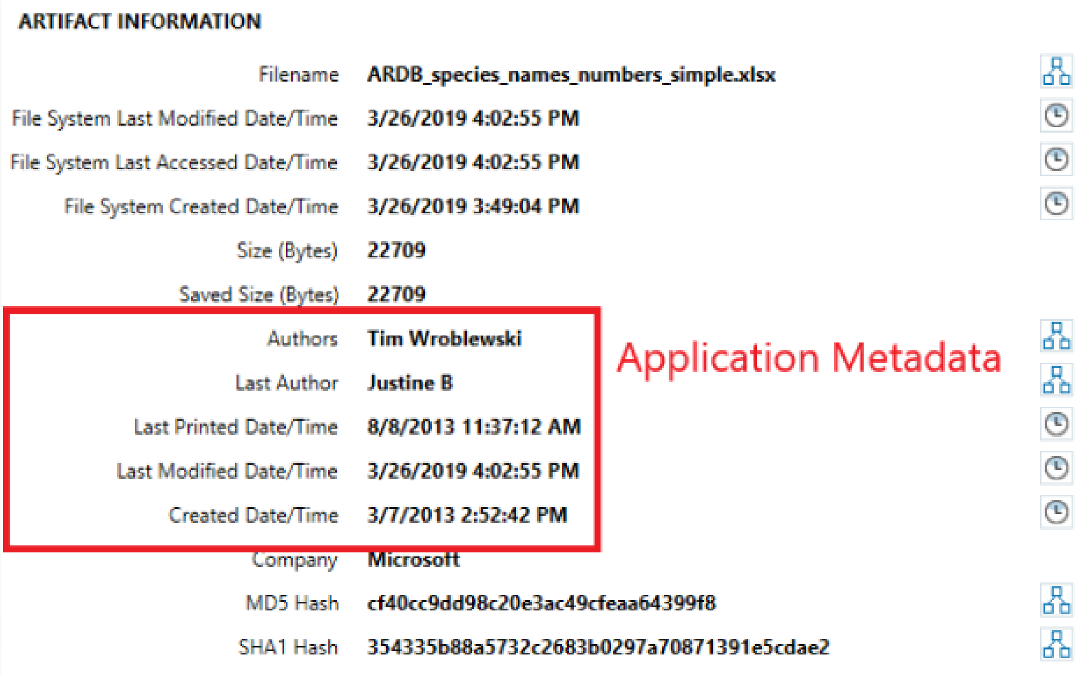
The Contents Pane includes a Details card, which lists some of the column data from the Evidence Pane. Included in this view will be additional metadata associated with the currently selected document. Document metadata (also known as Application Metadata) is typically stored internally (within the document or picture). This information is placed there by the application that was used to create the file. There is also File System Metadata. This metadata is typically stored and managed by the file system, primarily in the Master File Table ($MFT). Often there are time stamps stored internally as part of the metadata; this can sometimes result in inconsistencies between these times and the ones recorded by the file system. These inconsistencies should not be looked at as bad. For instance, if a document was modified, saved on a laptop and then printed, and the following day it was copied to a desktop computer, it would show a file-system-created time on the desktop computer of the more recent time, however internally the metadata times would reflect when it was modified and saved on the laptop and printed. The metadata times that are stored internally travel with the document whereas the file system times only stay with the document on that specific system. The following Application Metadata columns will be available in the Column View and in the Details card for most documents:

Component	Description
Title	The document title
Subject	The subject of the document
Authors	The authors / creators of the document
Keywords	The keywords contained in the document for (searching purposes)
Comments	The comments added to the document
Last Author	The last author to edit the document
Last Printed Date/Time (UTC)	The date and time the document was last printed
Last Modified Date/Time (UTC)	The last date and time the document was modified

Created Data/Time (UTC)	The data and time the document was created
Company	The company associated with the document

RTF and Text documents will not have Application Metadata. PDF files can store metadata similar to Microsoft Office documents. The Handgul Word Processor documents will have similar metadata to Office documents but include a few unique fields. See the Artifact Reference Guide for specific information.

Exercise Question 1
Open case file "**DFI_Practical_1_Case**". Go to "**Artifacts**" then open "**Documents**"
For the document "**ARDB_species_names_numbers_simple.xls**", note the differences in the date and times stamps for the Application Metadata and the File System Metadata. Application Metadata indicates that the document was created (came into existence) on 3/7/2013 by the author Tim Wroblewski. It was then last modified on 3/26/2019 by author Justine B. File System Metadata indicates that the file was created on 3/26/2019. This indicates the date that the file was first saved on the hard drive on which it is currently residing.

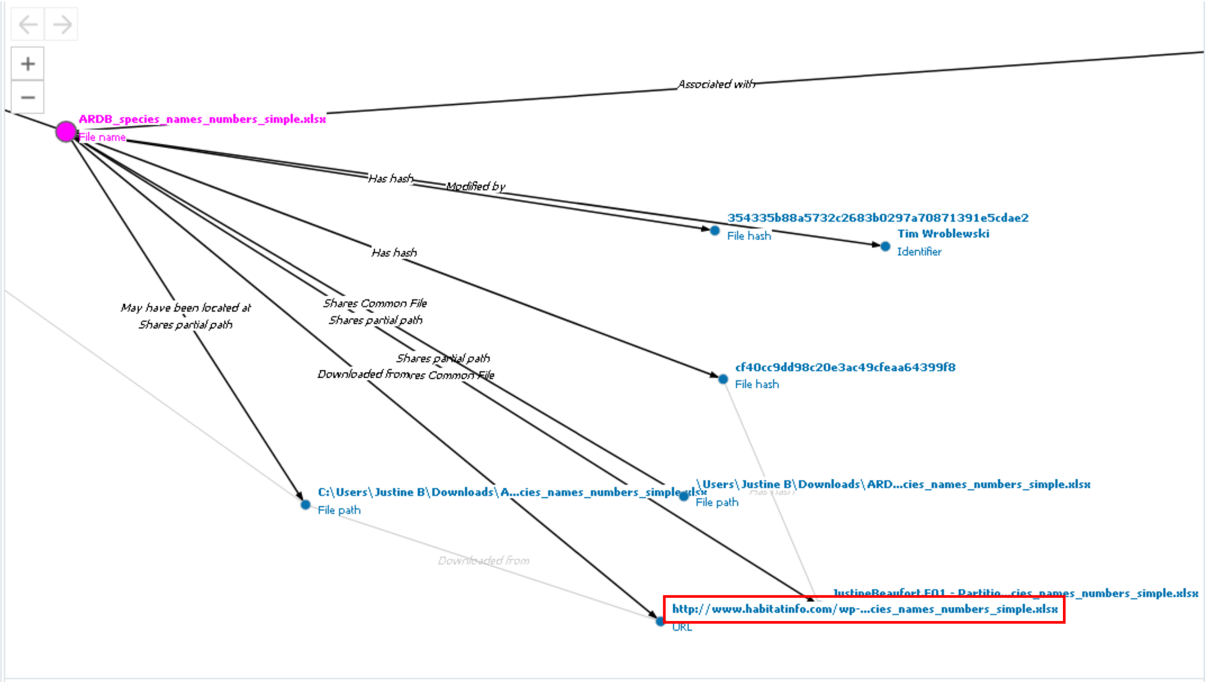


ARTIFACT INFORMATION

Filename	ARDB_species_names_numbers_simple.xlsx
File System Last Modified Date/Time	3/26/2019 4:02:55 PM
File System Last Accessed Date/Time	3/26/2019 4:02:55 PM
File System Created Date/Time	3/26/2019 3:49:04 PM
Size (Bytes)	22709
Saved Size (Bytes)	22709
Authors	Tim Wroblewski
Last Author	Justine B
Last Printed Date/Time	8/8/2013 11:37:12 AM
Last Modified Date/Time	3/26/2019 4:02:55 PM
Created Date/Time	3/7/2013 2:52:42 PM
Company	Microsoft
MD5 Hash	cf40cc9dd98c20e3ac49cfeaa64399f8
SHA1 Hash	354335b88a5732c2683b0297a70871391e5cdae2

Figure 3-3-5: Application Metadata

Based on the File System Metadata and the Application Metadata, we can conclude that the document was not created on the computer on which it currently resides. We can conclude that it was saved to the computer on 3/26/2019. Can we determine where the user obtained the file from? Hint: Click on the Connections icon next to the file name.

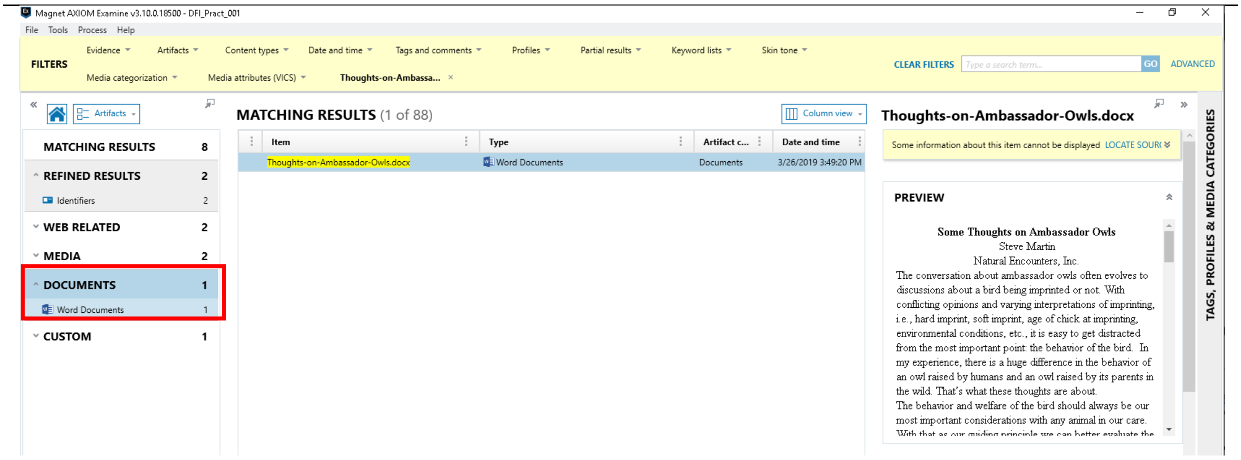

Associated with
ARDB_species_names_numbers_simple.xlsx
File name
Has hash
Modified by
354335b88a5732c2683b0297a70871391e5cdae2
File hash
Tim Wroblewski
Identifier
Has hash
May have been located at
Shares partial path
Shares Common File
Shares partial path
Shares partial path
Downloaded fromres Common File
cf40cc9dd98c20e3ac49cfeaa64399f8
File hash
C:\Users\Justine B\Downloads\A...cies_names_numbers_simple.xlsx
File path
\Users\Justine B\Downloads\ARD...cies_names_numbers_simple.xlsx
File path
Downloaded from
JustineBeaufort.E01 - Partitio...cies_names_numbers_simple.xlsx
http://www.habitatinfo.com/wp-...cies_names_numbers_simple.xlsx
URL

Exercise 2. Using Connections Explorer For Documents

Exercise Question 2

1. Go to Word documents and search for a document named “**Thoughts-on-Ambassador-Owls.docx**”.

How many results are there?



Magnet AXIOM Examine v3.10.0.18500 - DFI_Pract_001

MATCHING RESULTS	8
REFINED RESULTS	2
Identifiers	2
WEB RELATED	2
MEDIA	2
DOCUMENTS	1
Word Documents	1
CUSTOM	1

MATCHING RESULTS (1 of 88)

Item	Type	Artifact c...	Date and time
Thoughts-on-Ambassador-Owls.docx	Word Documents	Documents	3/26/2019 3:49:20 PM

Thoughts-on-Ambassador-Owls.docx

Some information about this item cannot be displayed LOCATE SOUR(

PREVIEW

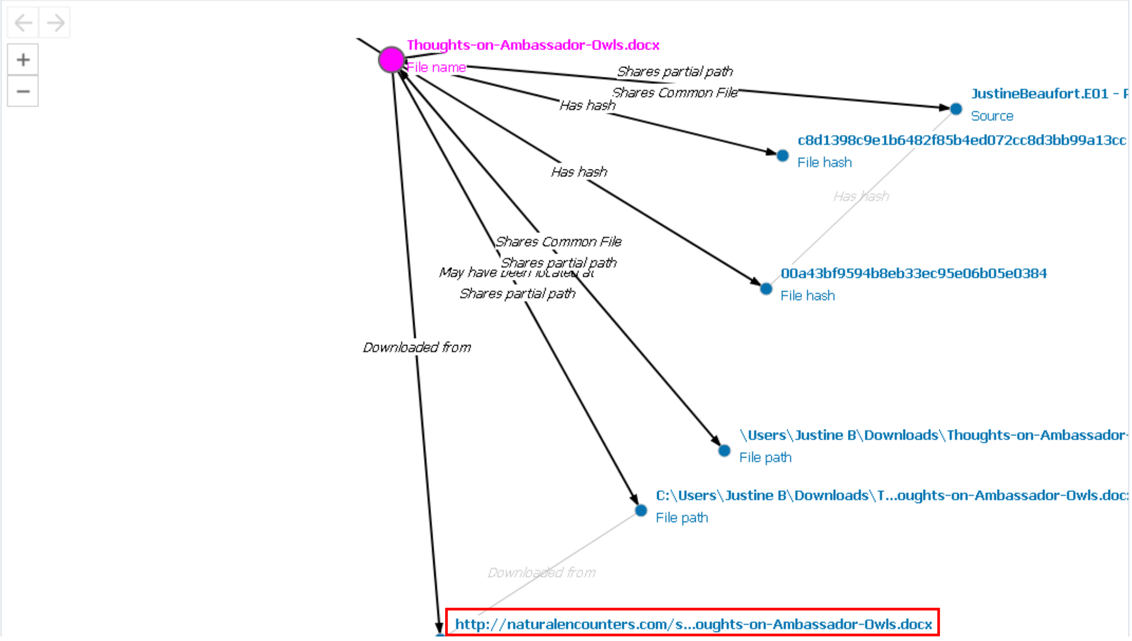
Some Thoughts on Ambassador Owls
Steve Martin
Natural Encounters, Inc.

The conversation about ambassador owls often evolves to discussions about a bird being imprinted or not. With conflicting opinions and varying interpretations of imprinting, i.e., hard imprint, soft imprint, age of chick at imprinting, environmental conditions, etc., it is easy to get distracted from the most important point: the behavior of the bird. In my experience, there is a huge difference in the behavior of an owl raised by humans and an owl raised by its parents in the wild. That’s what these thoughts are about.

The behavior and welfare of the bird should always be our most important considerations with any animal in our care. With that as our guiding principle we can better evaluate the

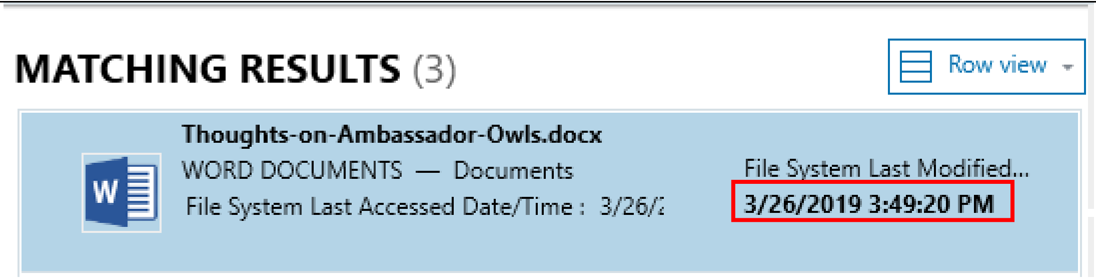
2. Click on the connections icon next to the file name Thoughts-on-Ambassador-Owls.docx.

What URL was this file downloaded from?



Thoughts-on-Ambassador-Owls.docx
File name
Shares partial path
Shares Common File
Has hash
JustineBeaufort.E01 - P
Source
c8d1398c9e1b6482f85b4ed072cc8d3bb99a13cc
File hash
Has hash
Has hash
Shares Common File
Shares partial path
May have been located at
Shares partial path
00a43bf9594b8eb33ec95e06b05e0384
File hash
Downloaded from
\Users\Justine B\Downloads\Thoughts-on-Ambassador-
File path
C:\Users\Justine B\Downloads\T...oughts-on-Ambassador-Owls.doc:
File path
Downloaded from
http://naturalencounters.com/s...oughts-on-Ambassador-Owls.docx

What was the date and time of the download?

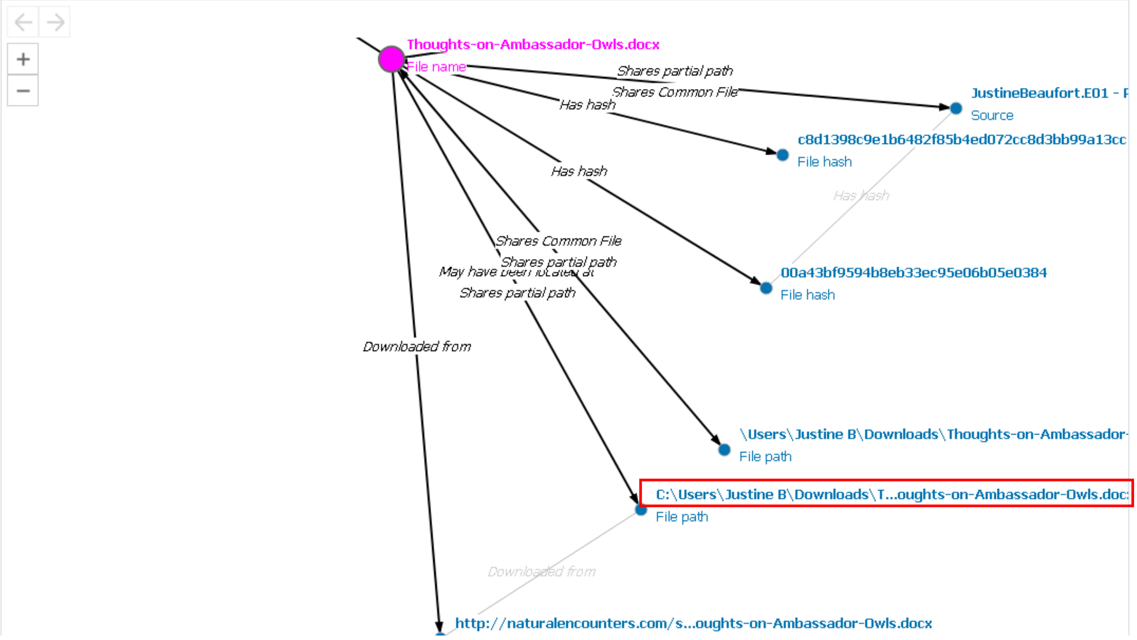


MATCHING RESULTS (3)

Thoughts-on-Ambassador-Owls.docx
WORD DOCUMENTS — Documents
File System Last Accessed Date/Time : 3/26/2

File System Last Modified...
3/26/2019 3:49:20 PM

Where was this document saved on the computer?



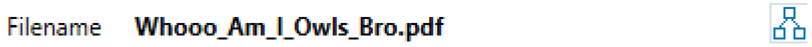
Thoughts-on-Ambassador-Owls.docx
File name
Shares partial path
Shares Common File
Has hash
JustineBeaufort.E01 - P
Source
c8d1398c9e1b6482f85b4ed072cc8d3bb99a13cc
File hash
Has hash
Has hash
Shares Common File
Shares partial path
May have been located at
Shares partial path
00a43bf9594b8eb33ec95e06b05e0384
File hash
Downloaded from
\Users\Justine B\Downloads\Thoughts-on-Ambassador-
File path
C:\Users\Justine B\Downloads\T...oughts-on-Ambassador-Owls.doc
File path
Downloaded from
http://naturalencounters.com/s...oughts-on-Ambassador-Owls.docx

Exercise 3. Using Connections Explorer For Documents

3. We want to tell the story about the document Whooo_Am_I_Owls_Bro.pdf.

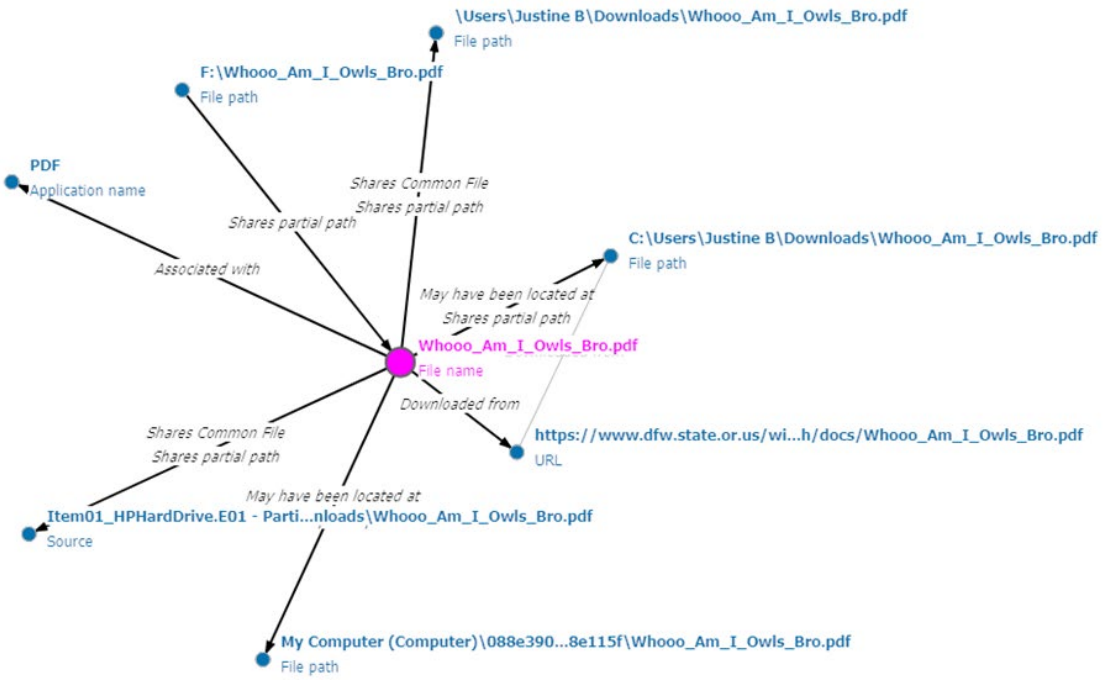
Exercise Question 3

4. Locate the document "**Whooo_AM_I_Owls_Bro.pdf**". Go to the Connections Explorer for the Filename.



Filename **Whooo_Am_I_Owls_Bro.pdf**

5. What story does this connections graph tell?


\Users\Justine B\Downloads\Whooo_Am_I_Owls_Bro.pdf
File path
F:\Whooo_Am_I_Owls_Bro.pdf
File path
PDF
Application name
Shares Common File
Shares partial path
Shares partial path
Associated with
C:\Users\Justine B\Downloads\Whooo_Am_I_Owls_Bro.pdf
File path
May have been located at
Shares partial path
Whooo_Am_I_Owls_Bro.pdf
File name
Downloaded from
https://www.dfw.state.or.us/wi...h/docs/Whooo_Am_I_Owls_Bro.pdf
URL
Shares Common File
Shares partial path
May have been located at
Item01_HPHardDrive.E01 - Parti...nloads\Whooo_Am_I_Owls_Bro.pdf
Source
My Computer (Computer)\088e390...8e115f\Whooo_Am_I_Owls_Bro.pdf
File path

Where was the document downloaded from?

What was used to download it?

When was it downloaded

Where was it saved to? If it was saved to more than one location, list all locations.

-- End --