A SIMPLE EXPERIMENT WITH MICROSOFT OFFICE 2010 AND WINDOWS 7 UTILIZING DIGITAL FORENSIC METHODOLOGY

Gregory H. Carlton
California State Polytechnic University
ghcarlton@csupomona.edu

ABSTRACT

Digital forensic examiners are tasked with retrieving data from digital storage devices, and frequently these examiners are expected to explain the circumstances that led to the data being in its current state. Through written reports or verbal, expert testimony delivered in court, digital forensic examiners are expected to describe whether data have been altered, and if so, then to what extent have data been altered. Addressing these expectations results from opinions digital forensic examiners reach concerning their understanding of electronic storage and retrieval methods. The credibility of these opinions evolves from the scientific basis from which they are drawn using forensic methodology.

Digital forensic methodology, being a scientific process, is derived from observations and repeatable findings in controlled environments. Furthermore, scientific research methods have established that causal conclusions can be drawn only when observed in controlled experiments. With this in mind, it seems beneficial that digital forensic examiners have a library of experiments from which they can perform, observe results, and derive conclusions. After having conducted an experiment on a specific topic, a digital forensic examiner will be in a better position to express with confidence the state of the current data and perhaps the conditions that led to its current state.

This study provides a simple experiment using the contemporary versions of the most widely used software applications running on the most commonly installed operation system. Here, using the Microsoft Office 2010 applications, a simple Word document, an Excel spreadsheet, a PowerPoint presentation, and an Access database are created and then modified. A forensic analysis is performed to determine the extent in which the changes to the data are identified. The value in this study is not that it yields new forensic analysis techniques, but rather that it illustrates a methodology that other digital forensic examiners can apply to develop experiments representing their specific data challenges.

1. INTRODUCTION

Digital forensic examiners are chartered with analyzing data from electronic devices with the goal of identifying current and previously stored data, and if possible, identifying the tasks that were performed causing the data to be in its current and previous states (Nelson, Phillips, & Steuart, 2010). When analyzing data, digital forensic examiners utilize forensically sound, scientific methods (Volonino, Anzaldua, & Godwin, 2007). Current scientific methodology establishes that experiments are necessary to identify causal conditions (Hoyle, Harris, & Judd, 2002).

This paper documents the development of a simple experiment involving the contemporary version of the most commonly used software application suite (i.e., Microsoft Office 2010) running on the most widely used operating system (i.e., Windows 7) at this time and stored on a USB flash drive, a ubiquitous storage medium. This storage medium was selected for this experiment, as these devices are relatively inexpensive, readily available, and digital forensic examiners should understand to what extent these devices differ from internal disk drives in the way in which data are stored.

It is very plausible that digital forensic examiners will find themselves in situations where they are presented with a USB flash drive as the sole piece of evidence in a legal matter or as a piece of evidence in a legal matter consisting of multiple storage devices. It is also very plausible that data contained within a USB flash drive obtained as evidence would include Microsoft Office documents.

Software vendors update their versions of operating systems and application software periodically, and each of these updates bring the potential for changes to the ways in which data are stored and retrieved. "Microsoft Office Word 2007 and Office PowerPoint 2007 use XML-based file formats as their default file format, and Microsoft Excel 2010 uses a newer binary format" (Microsoft Corporation, 2013). From the perspective of a digital forensic examiner, a consequence of the change to binary formats in Microsoft Office documents means that data contained within documents are no longer easily identifiable by keyword searches.

With data now stored within a different format that yields different search results, to remain current, a digital forensic examiner must take steps to understand the new format. There are various methods for digital forensic examiners to remain current pertaining to this matter, including training, researching the literature on this topic, or conducting first-hand experiments. As an extension to the scientific process, I suggest that it is prudent for digital forensic examiners to conduct a series of experiments to ensure that they understand and can explain the data contained within the electronic storage devices they encounter.

In the following pages, I will present a simple experiment to forensically analyze data stored on a USB flash drive. The data were stored using contemporary applications running on the most frequently used operating system, and simple modifications were made to the data.

These experiments are not provided to teach digital forensic methodology, but rather they are presented as a simple basis from which other scientists (i.e., digital forensic examiners and researchers) may expand upon to create a library of relevant experiments to further the science and benefit the profession of digital forensic examiners.

2. METHODOLOGY

The concept presented here is based on a scenario in which a digital forensic examiner obtains a single item of evidence, a USB flash drive, and the data on the device appear to be in the format of MS Office applications. Proper execution of the experiment requires that it is conducted in a controlled environment, and the overall structure of the experiment consists of four phases. The first phase is to begin with a known item, such as a freshly wiped and formatted USB flash drive. In the second phase, the experimenter will create known content onto this device and then obtain a forensic image of the device to serve as the control set for the experiment. The third phase of the experiment is to make specific alterations to the data on the USB flash drive and then obtain a forensic image of the device to serve as the dependent variable data set. The fourth phase of the experiment is to forensically analyze the dependent variable data set to measure the extent to which data contained within the control set are visible within the dependent variable data set. Each of these phases is expanded upon below.

2.1 Phase One – Storage Media

In the first phase of the experiment, it is essential to begin with a storage device in which its contents are known with certainty. There must be no opportunity for existing data on the device to contaminate the future findings when an analysis is performed. Initially, when the author envisioned this experiment, it seemed plausible that placing data on a newly purchased USB flash drive, without wiping the drive first, would make the conditions more realistic, as it was reasoned that users typically use these devices after purchasing them without wiping them; however, internal validity issues quickly surfaced from this approach (Babbie, 2004).

These internal validity issues arose when existing data were identified residing on the newly purchased device prior to adding any of the content data from phase two. Upon purchasing a 2GB Centon DataStick Pro USB flash drive, I first connected it to a Tableau USB write-blocking device, and created a forensic image, named "New Centon USB," of the device using EnCase Version 6. Using EnCase to search the New Centron USB image for non-null characters, I found

that much of the formatted disk volume within the storage device, excluding the volume boot record and file allocation table, contained non-null characters, thus posing a risk of data contamination in the future data analysis conducted in phase four. To remove this risk of contamination from residual data, I connected the USB flash drive directly to a USB port on my forensic workstation, and using EnCase, I wiped the newly purchased, USB flash drive, and formatted it as a FAT 32 volume with a volume label of OFFICETEST. FAT 32 was selected as the format, as this was the original format of the newly purchased USB flash drive, and by maintaining its original format, the outcomes will more closely match those that other examiners will encounter when analyzing these ubiquitous storage devices. After formatting the USB flash drive, I safely ejected the device, and I reconnected it the Tableau write-blocking device attached to a forensic workstation running EnCase.

With the wiped and formatted USB flash drive attached to the write-blocking device, I acquired a forensic image of the device, named "Wiped Office Test," and then analyzed it by searching for non-null characters. The results indicated that no residual data existed within the formatted volume of the device, excluding areas, such as the volume boot record and file allocation table. Now that we have determined that there are no residual data to contaminate future analyses on this device, we can conclude that the device is suitable for the experimental data and proceed to the second phase of this experiment.

2.2 Phase Two – Control Set Data

In the second phase of the experiment, the initial data were stored on the USB flash drive. The general concept here is to create simple content that is fully documented by the experimenter. It is essential that the contents of the initial data placed onto the USB flash drive are known. Care must be exercised to ensure that data entered are precisely documented. Four folders were created representing the applications used, namely Access, Excel, PowerPoint, and Word. Each application was used to create a corresponding document that was stored in the appropriate folder. Specific details of the data are described in Section 3. Data Descriptions.

After the initial data were stored on the USB flash drive, the storage device was safely ejected from the Windows 7 workstation, attached to a Tableau USB write-blocking device, and was forensically imaged and named "Initial Data" using EnCase Version 6. Once the Initial Data image was completed I proceeded to the third phase of the experiment.

2.3 Phase Three – Data Modifications

In this phase of the experiment, I reinserted the USB flash drive into the Windows 7 workstation, opened Windows Explorer, navigated to each of the application folders contained within the USB flash drive, and opened each

document by double-clicking on the document name to launch the appropriate application program while ensuring that the current folder was the one in which the document was stored.

Within each application, I made specific modifications to the data, as described in Section 3. Data Descriptions. Upon completing the data modifications, I clicked on the "save" function within each application to save the modified file under the same name as the original file. The rationale for this is to attempt to emulate the process typical users are likely to perform when modifying documents stored on USB flash drives. From a digital forensic examiner's prospective, I would expect that this process would result in the Windows 7's file management system creating a temporary file within the current folder, and when the save operation is performed, the Windows 7's file management system would delete the initial document file, and then rename the temporary file to the name of the initial document file.

If this process were to occur as described above, then a digital forensic examiner would expect to find, upon examining an image of the modified data, both the intact document file containing the modifications and the deleted document file containing the initial, unmodified data. These expectations thus become the null hypothesis for this experiment. While these expectations might seem reasonable to digital forensic examiners, given that versions of operating systems and application change, this experiment is necessary to validate those expectations. As the information presented in Section 4. Summary of Findings show, the results differ among the applications within the Office 2010 application suite.

Once the data modifications for each of the four application documents were completed, I safely ejected the USB flash drive from the Windows 7 workstation, connected the USB flash drive to the Tableau USB write-blocking device, and acquired a forensic image, named "After Modifications," using EnCase Version 6. Upon the completion of the data acquisition, I then proceeded to the fourth phase of the experiment to analyze the dependent variable data set, the "After Modifications" image file.

2.4 Phase Four – Analyzing the Dependent Variable Data Set

The fourth phase of this experiment consists of conducting a forensic analysis of the "After Modifications" image file, representing the dependent variable data set of the experiment. The steps taken during the analysis will likely vary from examiner to examiner, and the steps described within this paper are not intended to represent a comprehensive set of tasks. A thorough forensic analysis may consume a considerable amount of time, and that is beyond the scope of this paper. The focus here is to illustrate that tasks can be performed to recover the initial data, prior to the modifications, from the "After Modifications" image file, and these tasks yield varying degrees of success among the four applications of the Microsoft Office 2010 suite tested here.

To analyze the dependent variable data set, I opened the "After Modifications" image file within EnCase Version 6, and performed tasks including keyword searches, document viewing, and file exports. Details of the specific task operations performed are presented in Section 4. Summary of Findings.

After conducting a forensic analysis of the dependent variable set, a digital forensic examiner is then in a position to speak with a greater level of certainty regarding generalizing the state of data contained within evidence meeting similar conditions. This simple experiment design allows digital forensic examiners to easily modify parameters to meet data conditions to match their evidence, thus providing them with a flexible scientific resource.

3. DATA DESCRIPTIONS

Within this section the initial data and the modified data are described for each of the four Microsoft Office applications evaluated in this experiment. The data for the Microsoft Excel spreadsheet file is presented first followed by the data for the Microsoft PowerPoint presentation file, then the data for the Microsoft Word document file is discussed, and finally the data for the Microsoft Access database file is described. Within the discussion for each application, the initial data is presented first followed by the modified data. This section is limited to a description of the data, and the discussion on findings from the analysis of the data is presented in Section 4. Summary of Findings.

3.1 Microsoft Excel Data

A simple, one-worksheet, Excel spreadsheet file was created and subsequently modified in this experiment. The spreadsheet file named Excel 2010 Experiment.xlsx was stored on the USB flash drive within the Excel folder. After the file was created and saved with the initial data, it was opened, the data were modified, and it was saved under the same file name by clicking on the save icon within the Excel application (Shelly & Vermaat, 2011). It is interesting to note that when the file was subsequently opened for modification, Windows 7 created a temporary file named ~$Excel Experiment 2010.xlsx within the Excel folder of the USB flash drive, as shown in Figure 1 Excel Data Files. This temporary file serves the purpose as a placeholder for the edited data when the original file is opened. Should the edited file be closed without saving the changes, the original file remains intact, and the temporary file is deleted; however, if the user selects the save file function, the original file is deleted, and the temporary file containing the modified data is renamed to the name of the original file (Korfhage, 1997). The value of this method of deleting original data files and saving modified data files under the same name within the folder will become apparent during the subsequent data analysis, as discussed in Section 4. Summary of Findings.

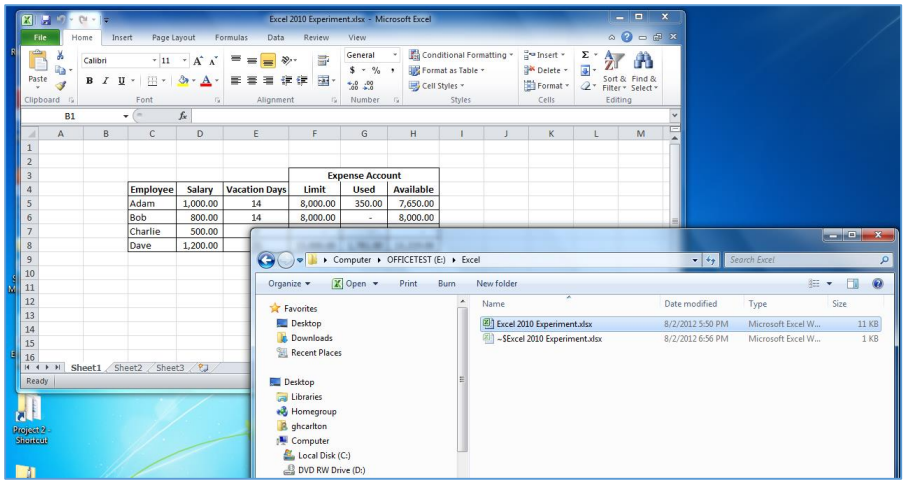

Figure 1 Excel Data Files

3.1.1 Initial Data

The initial data entered into the Excel spreadsheet file consisted of a worksheet containing four rows and six columns of data, plus column headings, as shown in Figure 2 Excel Initial Data.

			Expense Account		
Employee	**Salary**	**Vacation Days**	**Limit**	**Used**	**Available**
Adam	1,000.00	14	8,000.00	350.00	7,650.00
Bob	800.00	14	8,000.00	-	8,000.00
Charlie	500.00	7	-	-	-
Dave	1,200.00	21	15,000.00	1,781.00	13,219.00

Figure 2 Excel Initial Data

3.1.2 Modified Data

The initial data were modified to delete the 350.00 from the Expense Account Used column of the row for the employee Adam, thus resulting in a change in the Expense Account Available amount to 8,000.00. Changes to data in the row attributed to employee Bob consisted of a salary increase to 18,000.00, increase in vacation days to 21, and an increase in the Expense Account Limit column to 20,000, thus yielding an increase in the Expense Account Available column to 20,000. No changes were made to the row attributed to employee Charlie. The value in the Expense Account Used column for the row attributed to employee Dave was modified to 2,131.00, thus yielding an Expense Account Available value of 12,869.00. An example of the modified spreadsheet is shown in Figure 3 Excel Modified Data.

			Expense Account		
Employee	**Salary**	**Vacation Days**	**Limit**	**Used**	**Available**
Adam	1,000.00	14	8,000.00	-	8,000.00
Bob	18,000.00	21	20,000.00	-	20,000.00
Charlie	500.00	7	-	-	-
Dave	1,200.00	21	15,000.00	2,131.00	12,869.00

Figure 3 Excel Modified Data

3.2 Microsoft PowerPoint Data

A relatively simple PowerPoint presentation file named PowerPoint 2010 Experiment.pptx and consisting of three slides was created and subsequently modified in this experiment. The presentation file was stored on the USB flash drive within the PowerPoint folder. After the file was created and saved with the initial data, it was opened, the data were modified, and the file was saved under the same file name by clicking on the save icon within the PowerPoint application (Shelly & Vermaat, 2011).

3.2.1 Initial Data

The initial data entered into the PowerPoint 2010 Experiment.pptx presentation file consisted of three slides. The first slide contained two rows of text centered within the slide. The first row of text contained the words, "Newly Created PowerPoint," and the second row of text stated, "Prior to any modifications." The second slide consisted of a centered title stating, "Newly Created PowerPoint," followed by four bulleted, left justified text items that contained the following words:

- This is the first bullet item
- This is the second bullet item
- This is the third bullet item
- This is the fourth and last bullet item

The third slide in the initial data PowerPoint presentation file contained a centered title stating, "Newly Created PowerPoint," and centered within the slide was an inserted graphic image depicting a drawing of a human hand knocking on a door. A screenshot of the contents of the initial data version of the PowerPoint presentation file is shown in Figure 4 PowerPoint Initial Data.

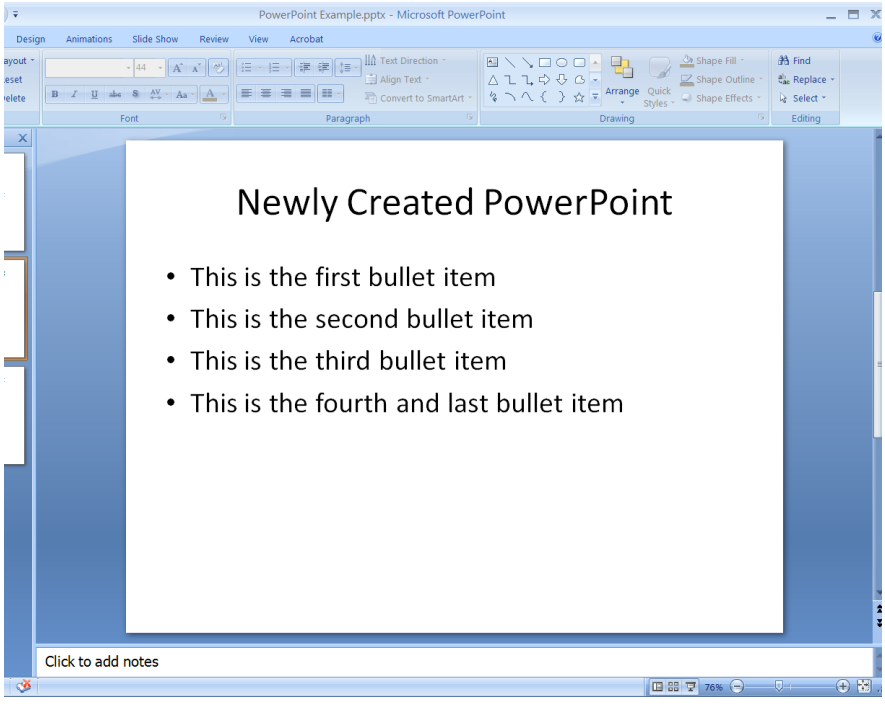

Figure 4 PowerPoint Initial Data

3.2.2 Modified Data

The initial data in the PowerPoint presentation file were modified with changes to each of the three slides. The second line of text on the first slide was modified to state, "After modifications." The bullet items on the second slide were modified, as shown below:

- This is the 1st bullet item
- This is the 2nd bullet item
- This is the fourth and last bullet item
- This is the newly added last bullet item

Additionally, the initial graphic image contained within the third slide was deleted and replaced by another graphic image. The newly inserted graphic image consists of a photograph of students lounging on a grassy area on a university campus. A screenshot of the contents of the modified data version of the PowerPoint presentation file is shown in Figure 5 PowerPoint Modified Data.

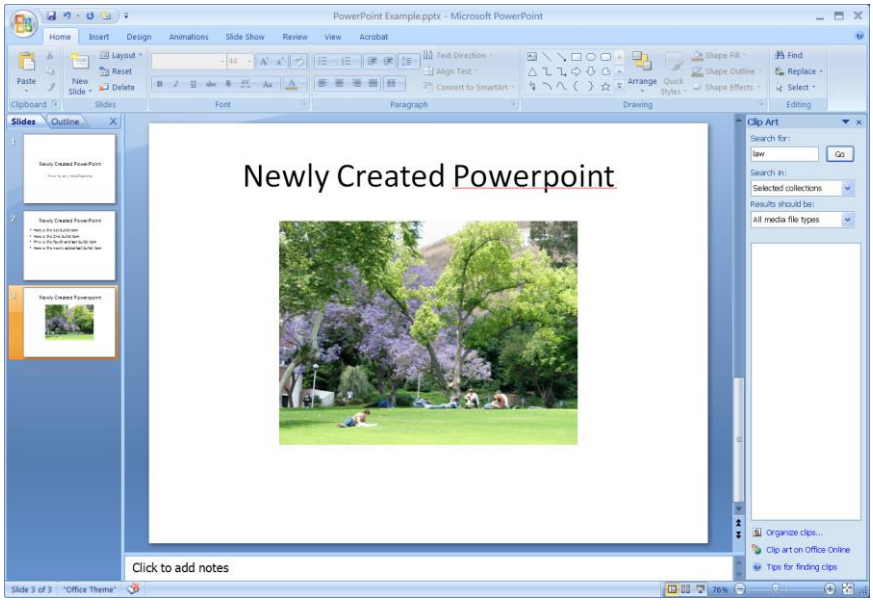

Figure 5 PowerPoint Modified Data

Similar to the temporary file created when the Excel 2010 Experiment.xlsx file was opened, Windows 7 also created a temporary file named ~$PowerPoint 2010 Experiment.pptx within the PowerPoint folder when the presentation file was opened, as shown in Figure 6 PowerPoint Data Files.

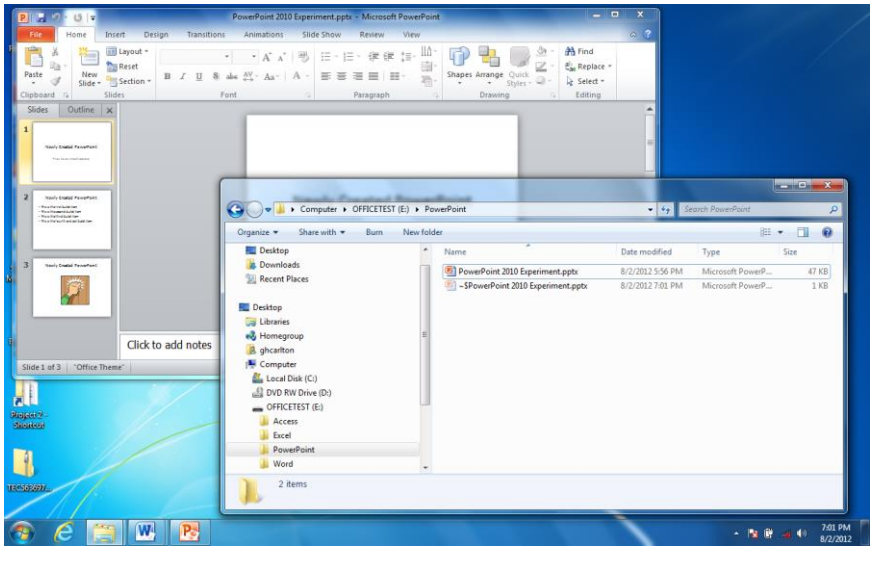

Figure 6 PowerPoint Data Files

3.3 Microsoft Word Data

3.3.1 Initial Data

The initial data entered into the Microsoft Word document consisted of three one-sentence paragraphs followed by an inserted graphic image showing a screenshot of a dialog box containing the completed status of the data acquisition of the "New Centon USB" image file from EnCase.

The three paragraphs contained the following text:

> This is the first paragraph of the initial, unmodified version of the Microsoft Office 2010 text used in this experiment.
>
> Changes will be made to this text, and the data from the two versions will be compared.
>
> Here is a screenshot of the data acquisition of the newly purchased USB thumb drive:

3.3.2 Modified Data

The modified data consisted of changes to the wording in each of the three one-sentences paragraphs contained in the initial data and the inserted graphic image was replaced with a screenshot of a dialog box containing the completed status of the data acquisition of the ""Wiped Office Test" image file from EnCase.

The three modified paragraphs contained the following text:

Here is the first paragraph of the Microsoft Word, version 2010, text used in this experiment.

> Additions, deletions, and modifications were made to this revised version of the text, and a forensic data analysis will be performed to analyze differences in the data.
>
> Here is a screenshot of the data acquisition of the wiped and formatted USB thumb drive:

It is noteworthy to observe, that unlike the temporary files created within the current folders when the Excel spreadsheet file or the PowerPoint presentation file were opened, Windows 7 did not create a temporary file within the Word folder of the USB flash drive. As shown in Figure 7 Word Data Files, although the folder options of Windows Explorer are set to show hidden files, the temporary file is not shown. This will play a significant role during the subsequent data analysis, and the author is concerned that many digital forensic examiners may not be aware of this condition.

To further substantiate the exceptional condition observed regarding a temporary file not being created within the current folder of the USB flash drive when a Microsoft Word 2010 document was opened under Windows 7, the author also

established a simple experiment to create a folder on the Windows 7 workstation's internal hard disk drive containing the bootable operating system. Opening a Microsoft Word 2010 document within the experimental folder on the internal hard disk did result in the creation of a temporary file similar to the results identified by opening Excel and PowerPoint files on the USB flash drive.

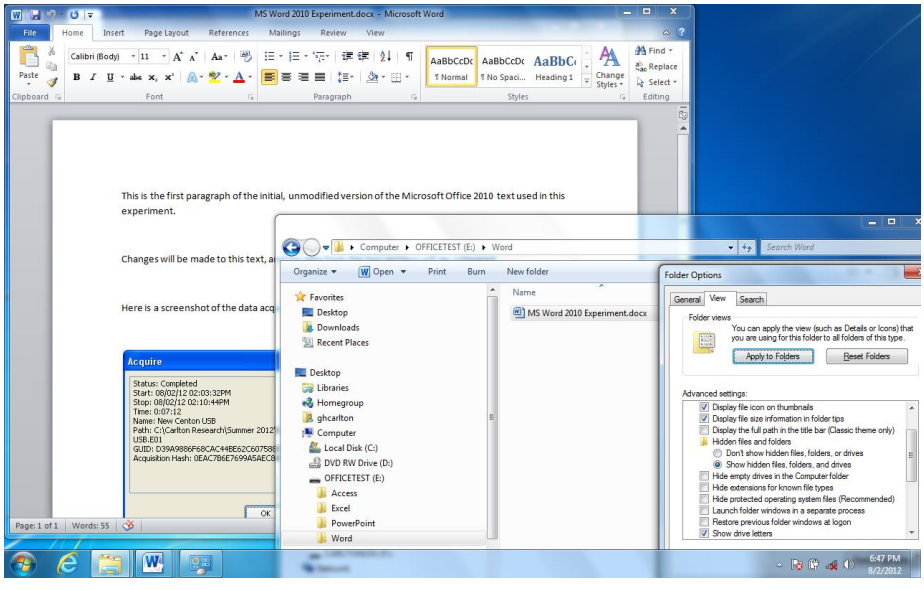

Figure 7 Word Data Files

3.4 Microsoft Access Data

A simple Microsoft Access database file consisting of a single table was created and subsequently modified in this experiment. The database file was stored on the USB flash drive within the Access folder. After the file was created and saved with the initial data, it was opened, the data were modified, and it was saved under the same file name by closing the table and exiting the Access application (Shelly & Vermaat, 2011).

3.4.1 Table Design

A single entity (i.e., table) named Employee Table, was created, consisting of eight attributes (i.e., columns). Six of the attributes correspond with the six columns described in Section 3.1 above within the Excel spreadsheet data, namely: Employee Name, Salary, Vacation Days, Expense Account Limit, Expense Account Used, and Expense Account Available. In addition to these six attributes, two additional attributes were included in the table design, Employee ID and Notes. The Employee ID and the Vacation Days attributes are of long integer type, Employee Name is text, Salary, Expense Account Limit, Expense Account Used, and Expense Account Available are currency, and the Notes

attribute is of type memo. A screenshot of the table design is shown in Figure 8 Access Entity Structure.

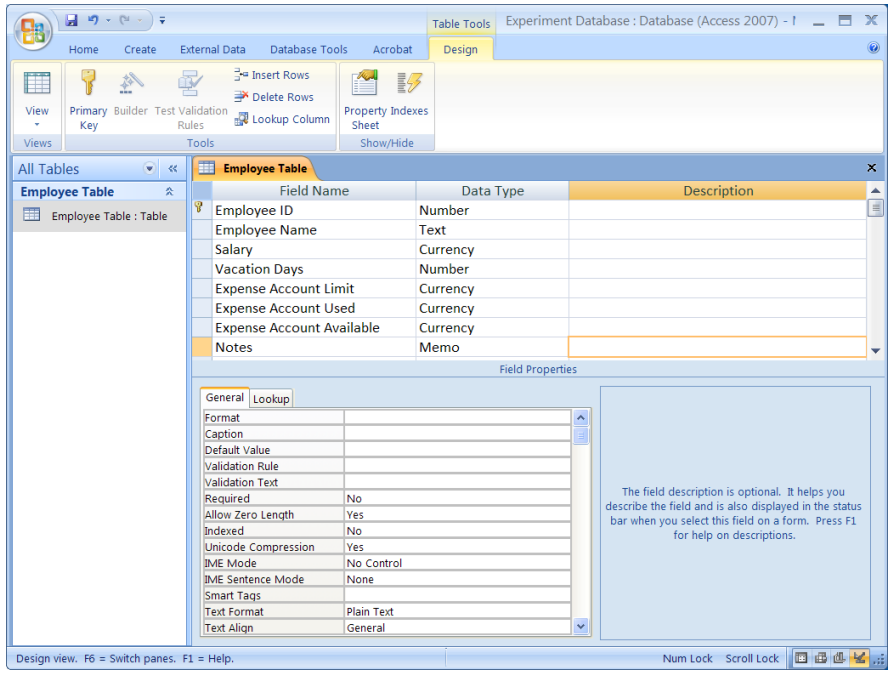

Figure 8 Access Entity Structure

3.4.2 Initial Data

The initial data entered into the Employee Table matches the initial data entered into the Excel spreadsheet, as describe in Section 3.1.1, with the additional values for the Employee ID and Notes attributes. The Employee ID values are entered as sequential integers, beginning with the value of 1 for the record corresponding with the Employee Name of Adam, and continuing to the value of 4 for record containing the value of Dave in the Employee Name attribute. The Notes attributes contains text stating, "Here are the initial notes regarding," followed by the name of the value contained in the Employee Name attribute (e.g., "Here are the initial notes regarding Adam"). A screenshot of the initial data is shown in Figure 9 Access Initial Data.

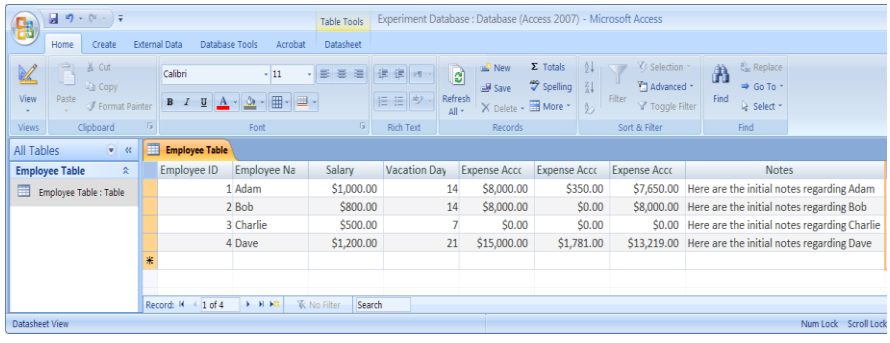

Figure 9 Access Initial Data

3.4.3 Modified Data

The modified Access database consisted of changes identical to those made in the modified Excel spreadsheet described in Section 3.1.2, with the following additional changes. Edits were made to the Notes attributes for each record to replace the word "initial" with the word "modified." Also the record with the Employee ID value of 3 and the Employee Name value of Charlie was deleted.

After the modifications were made, the database was closed, and the Access application was exited. The internal utility to compact a database within the Access application was not performed, thus allowing a better opportunity to reclaim deleted database records in the subsequent forensic analysis (Shelly & Vermaat, 2011).

The file management system of Windows 7 treats opening an Access database file differently than it does other Microsoft Office files, such as Excel, PowerPoint, or even Word. Instead of creating a temporary file within the current folder, a lock file is created. In this example, the file is named Access 2010 Experiment.laccdb. Notice that this is not a typical, hidden temporary file that begins with the "~$" prefix, but rather a non-hidden, temporary file that is deleted at the time the database file is closed. The lock file exists to facilitate record-level locking of database edits in a multi-user environment (Baeza-Yates & Ribeiro-Neto, 1999). Also, unlike the other office applications whereby temporary files are created to contain data edits, Access database files permit data modifications directly within the original database file. While this technique is advantageous in a large, multi-user database environment to ensure that data are updated immediately upon the completion of a transaction, this technique presents limitations on digital forensics examiners regarding the availability of historical data prior to modification (Elmasri & Navathe, 2003).

4. SUMMARY OF FINDINGS

After the data modifications were completed, as described in Section 3 of this report, the USB flash drive was safely ejected from the Windows 7 workstation and connected to a Tableau USB write-blocking device connected to a forensics workstation running Encase Version 6. A forensic image, named "After Modifications" was acquired, and a cursory forensic analysis was conducted to determine whether evidence of the initial data set (i.e., control set) were easily identifiable within the dependent variable data set of the experiment, the "After Modifications" image file.

Readers should be aware that it is not my intention to suggest that the tasks performed in this forensic analysis are thorough, complete, or the most efficient tasks to identify evidence items; instead, my intention is to suggest a simple, repeatable, experiment based on scientific research methods that practitioners of digital forensic analysis or other researchers can easily duplicate or modify to address specific data circumstances in which they encounter. There are a number of validated digital forensic analysis tools available to digital forensics examiners, and I chose to use EnCase Version 6 in this experiment due to its general popularity in the contemporary marketplace. In fact, the author is suggesting that others conduct more extensive analyses using a variety of tools on their respective "After Modifications" images. The point here is that care must be taken to ensure the original content of the data set is known, the modifications are also known, and the process is controlled in a forensically sound manner that adheres to scientific methods for experiments. Then after obtaining the dependent variable data set (e.g., After Modifications), the examiner is free to utilize any available tools to examine the dependent variable data set.

The specific tasks used to analyze each of the four Microsoft Office 2010 applications are presented below.

4.1 Summary of Microsoft Excel Experiment Findings

Using EnCase Version 6 to analyze the After Modifications image, selecting the Excel folder from the Navigation Pane reveals nine data items in the table pane. By inspecting each of these nine data items, it becomes apparent quickly that earlier versions of the Excel 2010 Experiment.xlsx file are retained within the data allocated to the Excel folder (Guidance Software, Inc., 2008). The "Doc" setting within the detail pane of EnCase presents the data in the format of the creating application; therefore, as illustrated in Figure 10 EnCase View of Excel data, the initial data (i.e., control set) are visible.

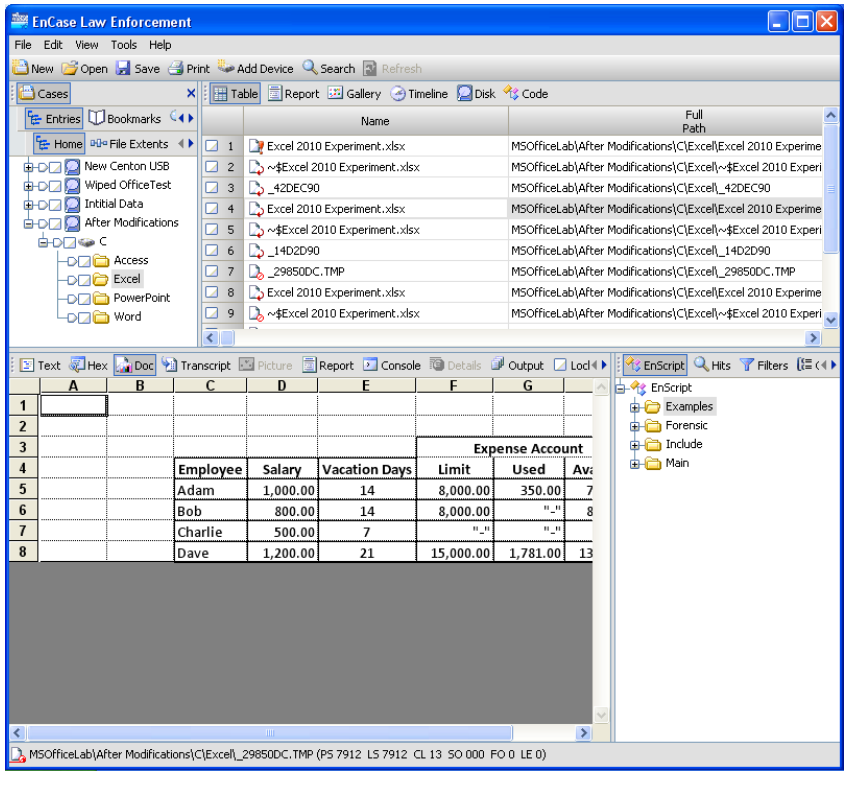

Figure 10 EnCase View of Excel data

4.2 Summary of Microsoft PowerPoint Experiment Findings

Similar to the results described in section 4.1 Summary of Microsoft Excel Experiment Findings, the results from analyzing the PowerPoint data yield evidence of the initial data. Here, using EnCase Version 6 to analyze the After Modifications image, selecting the PowerPoint folder from the Navigation Pane reveals data items in the table pane attributed to deleted versions of the PowerPoint 2010 Experiment.pptx file. Again, by inspecting each of these data items with the "Doc" setting in the EnCase view pane presents the data in the format of the creating application (Guidance Software, Inc., 2008). As illustrated in Figure 11 EnCase View of PowerPoint Data, the initial data (i.e., control set) are visible. In this example, the data presented within the EnCase view pane are displayed with three tabs positioned at the bottom the view pane window that represents each of the three slides contained within the presentation file. Clicking on these tabs, labeled Image 1, Image 2, and Image 3, results in EnCase displaying the contents of the corresponding slide (Guidance Software, Inc., 2008).

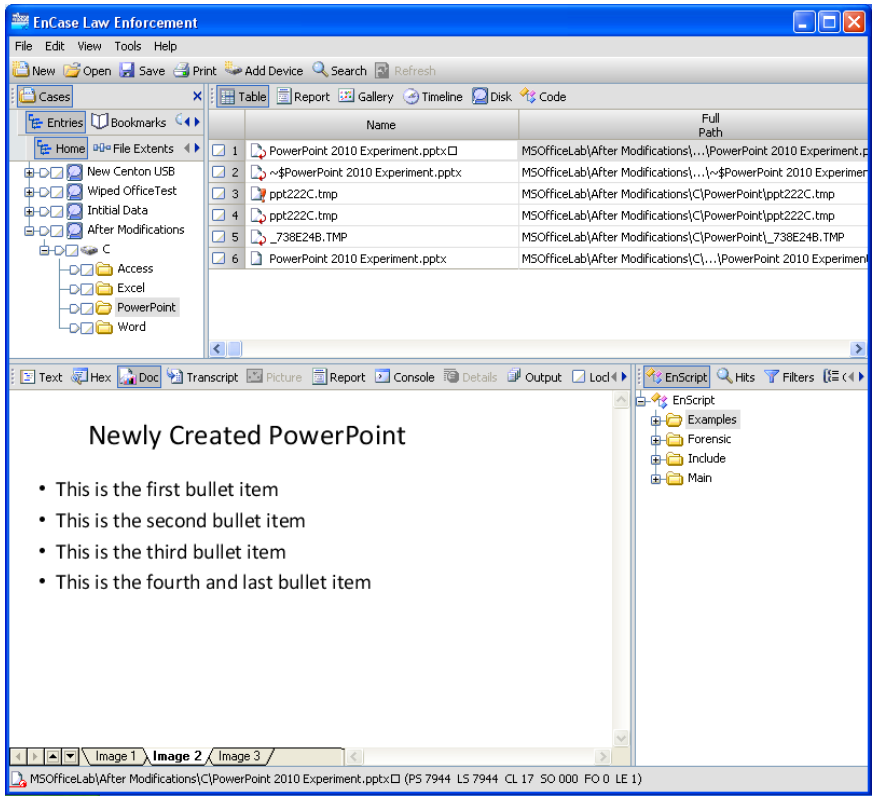

Figure 11 EnCase View of PowerPoint Data

4.3 Summary of Microsoft Word Experiment Findings

One might expect that the findings from an analysis of the Microsoft Word data would produce similar results as did the analyses of Excel and PowerPoint; however, there are distinct differences in the way in which the temporary files are handled.

As shown in Figure 12 EnCase Table of Word files, EnCase shows six data items listed within its table pane when the Word folder is selected from the After Modifications data image. However, examining these yields results in only one of the data items being visible in the view pane while using the Doc setting, and that document contains the modified data. No evidence items from the original data, prior to the modifications, are viewable in the native, document view (Guidance Software, Inc., 2008).

Additionally, keyword searches within EnCase were performed on five terms that were included in the original data and omitted from the modified data. These five keyword search terms were:

- unmodified (not case sensitive)

- This (case sensitive)
- first (case sensitive)
- changes will be made (not case sensitive)
- newly purchased (case sensitive)

No search hits were identified from a keyword search analysis of the After Modifications data image (Guidance Software, Inc., 2008). A visual inspection of the six data items listed within the table pane was also performed by viewing each data item in the text mode of the view pane within EnCase. This visual inspection also yielded negative results, as the data payload within the docx document file is not presented in the standard, ASCII format. Additionally, the six data items identified within the Encase table pane, as shown in Figure 12 EnCase Table of Word files were copied to the export folder of the forensic workstation, and an attempt was made to open each one of them with Microsoft Word. Only the after modifications version of the document file opened, and the initial data that had been removed were not visible.

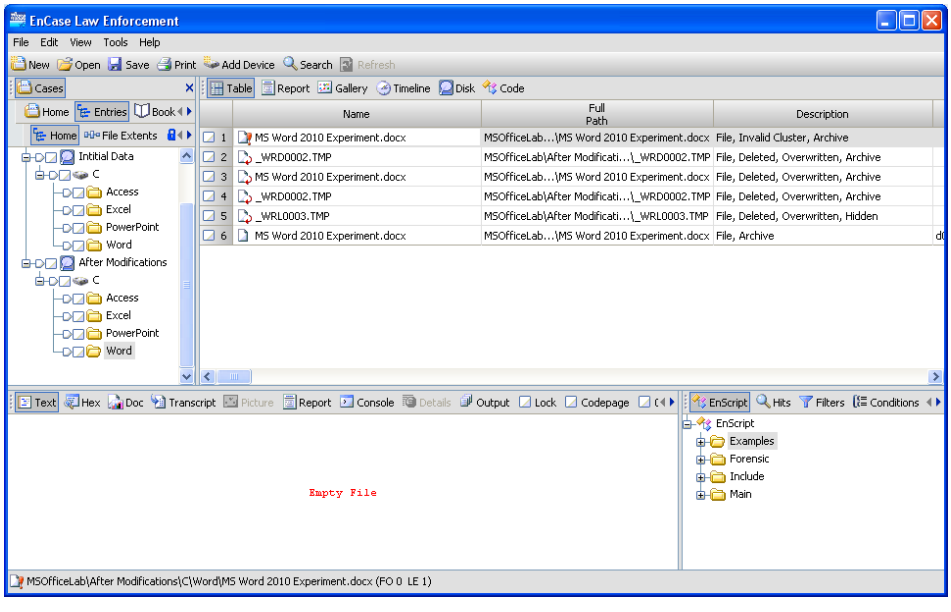

Figure 12 EnCase Table of Word files

4.4 Summary of Microsoft Access Experiment Findings

As shown in Figure 13 EnCase Table of Access Files, EnCase shows six data items listed within its table pane when the Access folder is selected from the After Modifications data image. However, none of these data items are viewable from the default file viewers available through the Doc tab of the view pane within EnCase Version 6. Manually analyzing the contents of the data items through the text or hex tabs within the view pane reveals that the data are stored

in a proprietary format instead of standard ASCII. Lastly, the data items were copied outside of the protected forensic environment of the .E01 file to an export folder on the forensic workstation, and I attempted to open each one of these exported files with Access 2010 (Guidance Software, Inc., 2008). Only the file containing the modified data opened, and none of the data from the initial data set that had been modified or deleted were visible.

Additional analyses were performed, including attempting to expand the .accdb files using the compound file function from EnCase's table pane and keyword searches for data unique to the original data set (Guidance Software, Inc., 2008). The results from these additional analyses were negative.

The author is not positing as assertion that the original data is not findable in a digital forensics analysis, but instead, it is not readily available from a simple, cursory analysis. This limitation is due to the proprietary format of the database, and this limitation can likely be overcome by obtaining one of many third-party software applications available within the marketplace that touts the ability to retrieve deleted data from Microsoft Access database files. One example of an application to recover data from within Microsoft Access database files is AccessFix by Cimaware (Cimaware, 2013).

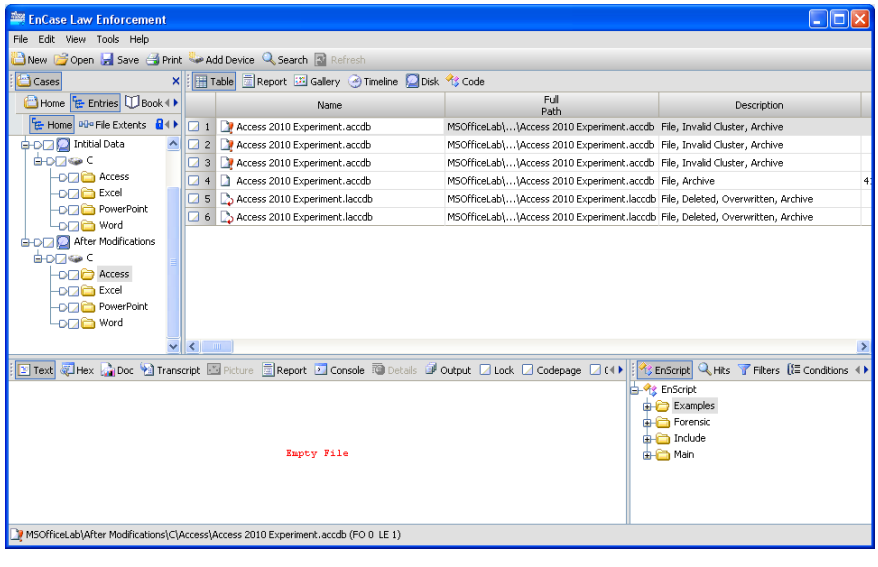

Figure 13 EnCase Table of Access Files

5. LIMITATIONS

Within this paper a simple set of experiments are presented whereby practitioners and researchers of digital forensics can use to perform observations of widely used applications in a controlled environment. These experiments are very limited in their scope, and they are not intended to be comprehensive, either

in their use of applications, treatment of data modifications, or in the analysis techniques presented. Rather than being a comprehensive set of experiments, these experiments should be viewed as templates from which other practitioners or researchers may alter to create experiments that match their specific data scenarios.

Limitations of this study include unique circumstances due to the version of Windows 7, the version of Microsoft Office 2010, the four Microsoft Office 2010 applications evaluated, the use of a USB flash drive as the data storage device, and version 6 of EnCase. Other application software, operating systems, storage media, and forensic analysis tools may yield different results.

Another limit of this study is based upon the presumption that digital forensics examiners should, on occasions, perform experiments to scientifically measure the treatment of data by operating systems and applications under specific storage media constraints. This presumption also relies on the unmeasured assumption that a substantial number of digital forensic examiners do not currently perform experiments in order to measure the affects of data modifications. This presumption leads directly to calls for additional research, as described in the following section.

6. CALL FOR ADDITIONAL RESEARCH

This study presents a collection of simple experiments from which digital forensics examiners can measure the extent to which they can identify original data stored on a USB flash drive that has subsequently been modified when using four widely used Microsoft 2010 applications under a Windows 7 operating system.

The concept for the need of these experiments was based on an anecdotal observation by the author, a practitioner, researcher, and instructor of digital forensics; however, this need should be measured through an empirical study. Two questions are immediately apparent that should be addressed, namely, is there a benefit from digital forensics examiners conducting experiments, and to what extent do digital forensic examiners perform experiments. Primary research, perhaps using grounded theory, will address these questions (Carlton, 2007).

Should an empirical study as described above yield the conclusion that there is both a benefit and a need for experiments, it would then be beneficial to develop a library of simple experiments from which digital forensic examiners may draw. The development of a library of experiments provides an excellent opportunity for academia to participate in the development of this body of knowledge. Lastly, if the usage of experiments is adopted by the digital forensic community of practitioners and researchers, then perhaps standard language concerning digital

forensic experiments can be appended to the work Cohen has suggested regarding a consensus in terminology (Cohen, 2011).

REFERENCES

Babbie, E. (2004). *The Practice of Social Research*, 10th ed. Belmont, CA: Thompson Learning.

Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern Information Retrieval.* New York, NY: ACM Press.

Carlton, G. H. (2007). A grounded theory approach to identifying and measuring forensic data acquisition tasks. *Journal of Digital Forensics, Security, and Law, 2* (1), 35-56.

Cimaware. (2013). *Repair Access database files with AccessFIX database recovery Software.* Retrieved from http://www.accessfix.com on June 10, 2013.

Cohen, F. (2011). Putting the science in digital forensics. *Journal of Digital Forensics, Security, and Law , 6* (1), 7-14.

Elmasri, R., & Navathe, S. B. (2003). *Fundamentals of Database Systems*, 4th ed. Boston, MA: Addison-Wesley.

Guidance Software, Inc. (2008). *EnCase Enterprise Version 6.10 User's Guide.* Pasadena, CA: Guidance Software, Inc.

Hoyle, R. H., Harris, M. J., & Judd, C. M. (2002). *Research Methods in Social Relations*, 7th ed. Boston, MA: Thompson Learning.

Korfhage, R. R., & Spencer, M. (ed). (1997). *Information Storage and Retrieval.* New York, NY: John Wiley & Sons, Inc.

Microsoft Corporation. (2013). *Understanding Office binary file formats.* Retrieved from http://msdn.microsoft.com/en-us/library/office/gg615407%28v=office.14%29.aspx on July 21, 2013.

Nelson, B., Phillips, A., & Steuart, C. (2010). *Guide to Computer Forensics and Investigations*, 4th ed. Boston, MA: Course Technology Cengage Learning.

Shelly, G. B., & Vermaat, M. E. (2011). *Microsoft Office 2010 Introductory.* Boston, MA: Course Technology Cengage Learning.

Volonino, L., Anzaldua, R., & Godwin, J. (2007). *Computer Forensics Principles and Practices.* Upper Saddle River, New Jersey: Pearson Prentice Hall.