2-2012

Cloud Computing and the Implications for E-Discovery

Ann H. Ziluck

APUS Library Capstone Submission Form

This capstone has been approved for submission to and review and publication by the APUS Library.

Student Name [Last, First, MI] *	Ziluck	Hearn	Ann
Course Number [e.g. INTL699] *	ITCC698	**Paper Date** [See Title pg.]	01/2012
Professor Name [Last, First] *	Watson-Stone, Novadean		
Program Name * See list	Master of Science in Information Technology		
Keywords [250 character max.]	E-Discovery, Cloud Computing, Digital Forensics		
Passed with Distinction * Y or N	Y		
Security Sensitive Information * Y or N	N		
IRB Review Required * Y or N	N	If YES, include IRB documents in submission attachments.	
Turnitin Check * Y or N	Y	All capstone papers must be checked via Turnitin.	

* Required

Capstone Approval Document

The thesis/capstone for the master's degree submitted by the student listed (above) under this title *

INFORMATION TECHNOLOGY - DIGITAL FORENSICS

has been read by the undersigned. It is hereby recommended for acceptance by the faculty with credit to the amount of 3 semester hours.

Program Representatives	Signatures	Date (mm/dd/yyyy)
Signed, 1st Reader * [capstone professor]		
Signed, 2nd Reader (if required by program)		
Recommendation accepted on behalf of the program director *	Novadean Watson-Stone — Digitally signed by Novadean Watson-Stone DN: cn=Novadean Watson-Stone, o=American Public University System, ou=APUS, email=nwatson@apus.edu, c=US Date: 2016.08.23 09:44:25 -04'00'	08/23/2016
Approved by academic dean *	Daniel L Welsch — Digitally signed by Daniel L Welsch DN: cn=Daniel L Welsch, o=American Public University System, ou=Academics, STEM, email=dwelsch@apus.edu, c=US Date: 2016.10.14 14:07:30 -04'00'	

* Required

AMERICAN PUBLIC UNIVERSITY SYSTEM

Charles Town, West Virginia

CLOUD COMPUTING AND THE IMPLICATIONS FOR E-DISCOVERY

A thesis submitted in partial fulfillment of the

requirements for the degree of

MASTER OF SCIENCE

in

INFORMATION TECHNOLOGY - DIGITAL FORENSICS

by

Ann Hearn Ziluck

Department Approval Date:
January 2012

DECLARATION

I hereby declare that the work has been done by myself and no portion of the work contained in this Thesis has been submitted in support of any application for any other degree or qualification on this or any other university or institution of learning.

_____

Ann Hearn Ziluck

ACKNOWLEDGEMENTS

I wish to thank my APUS graduate professors Dr. Karen Paullet, Dr. Diane Barrett, and Professor Robert Marlett who have encouraged me to strive for excellence and to reach for my academic goals. If not for their constant communications and support throughout my undergraduate degree, and the exciting discussions about network security, digital forensics and project management, I would not have immediately pursued a graduate degree. These dynamic professors continually 'threw down the academic gauntlet' forcing me to respond to their challenges. It seemed impossible to end the lively dialog with these three individuals after an undergraduate degree. Knowing they were all teaching in the newly launched Masters program in Information Technology Digital Forensics made it a very simple decision to continue my education in this exciting program. I would also like to thank Professors Sammy Abaza, and Murali Ramaswamy for serving as my technical mentors throughout my coursework. Their tireless support for a myriad of questions well outside the general course outlines was greatly appreciated.

Lastly, this document could not have been crafted without the continual support, patience and good humor of Dr. Novadean Watson-Stone. While I endeavored to write about the emerging technologies of cloud computing and e-discovery, she provided continual feedback, moral and academic support, editing guidance, and went to great lengths to procure requested sources that were in many cases only a few weeks old. Her guidance, enthusiasm and confidence in my abilities were invaluable.

DEDICATION

I dedicate this thesis to my son Michael Ziluck, my boyfriend Thomas Mondak, my parents John W. and Joan R. Hearn, my brother John W. Hearn, Jr. and my sister Barbara Allen-Lyall. Without their patience, understanding and support, academic encouragement, and most of all love, the completion of this work would not have been possible.

ABSTRACT

CLOUD COMPUTING AND THE IMPLICATIONS FOR E-DISCOVERY

by

Ann Hearn Ziluck

American Public University System, January XX, 2012

Charles Town, West Virginia

Dr. Novadean Watson-Stone, Thesis Professor

E-Discovery poses many challenges to legal and information technology teams who must quickly produce information required for litigation. The extraordinary volumes, types, electronic and manual formats, and repositories of data produced by even small organizations must be quickly searched and analyzed as part of the E-Discovery process. Data created and stored in cloud computing, social media and mobile computing environments intensify the typical complexities of e-discovery since traditional advanced search analytics, digital forensics and E-Discovery techniques are more difficult or in some cases impossible with these technologies. Cloud computing and social media environments often prevent extraction of metadata required for defensible production of information. Digital forensics is impossible in some cloud computing environments, preventing experts from obtaining defensible and reproducible data sets. Legal and information technology teams must carefully consider the use of cloud computing and social media environments and ensure their corporate constituents that they are empowered to perform structured, effective, and scientific e-discovery practices for regulatory and legal compliance.

Table of Contents

Cloud Computing - The Impact on E-Discovery

Introduction

E-discovery is an emerging and challenging field that combines the efforts of both legal and technology professionals in the production of electronic information required for litigation. Litigants are required to provide electronic information in a commonly agreed upon format often under extraordinary time constraints dictated by the courts. Organizations are required to search for and preserve documents and data, and must take reasonable steps to do so. This includes notifying all parties that might have the ability to destroy, alter, or otherwise compromise relevant information, and monitoring individual data custodians. The task of searching for and preserving the integrity of data is monumental, even when considering the volume of information created by a small organization that must search through all business transactions, paper documents, backup tapes and historical archives, foreign language documents, email and any Web 2.0 information (Schuler, Peter, & Vincze, 2009).

Cloud computing is an emerging paradigm that provides organizations with remote, scalable, on-demand network computing resources that require only minimal management effort. It also offers substantial cost savings over traditional datacenter network operations. However, Mell and Grance (2011) from the NIST caution readers that "security, interoperability, and portability" (p. 3) are some of the major considerations or barriers to widespread adoption of these methodologies. Brodkin (2008) from the Gartner Group agrees, and adds, "Data integrity, recovery, privacy and regulatory compliance are key issues to consider" (p. 1). Brodkin (2008) also makes a strong statement that "Investigating inappropriate or illegal activity may be impossible in cloud computing" (p. 2). Cloud hosting providers may not be willing or able to produce network logs due to shared tenancy on internationally dispersed and constantly changing

servers. Brodkin (2008) continues by quoting Gartner researchers Heiser and Nicolett (2008) who commented that if the cloud vendor cannot not contractually commit their support for specific investigations and e-discovery practices, and demonstrate previous successful instances of discovery compliance, then "the only safe assumption is that investigation and discovery requests will be impossible" (Brodkin, 2008, p. 1; Heiser & Nicholett, 2008, p. 4).

Rule 26 of the Federal Rules of Civil Procedure (FRCP) according to Cornell University Law School Staff (n.d.) are explicit with regard to the timely production of documents and information stored in electronic form owned by the disclosing party. Initial production and disclosure of information under the FRCP is generally required within 30 days of the Rule 26 conference between the two litigants. The literature suggests that cloud computing intensifies the common difficulties involved in e-discovery projects, taking them to an exponentially greater level, and may prevent discoverability entirely.

Problem Statement

Cloud computing environments can be used to host entire enterprise systems to empower a mobile workforce with the added benefits of considerably reducing costs and resources required to maintain a networked computing infrastructure. Cloud computing environments are also linked with social media and Web 2.0 data that plays an important part in the scope of discoverable enterprise data. Cloud computing seems to exacerbate the complexities of the practice of e-discovery making advanced search analytics, digital forensics and e-discovery far more difficult or impossible. Further research could illustrate these complexities and help to educate organizations when making cloud computing choices, and change the contractual assurances that cloud vendors must ultimately provide to empower their customers to perform e-discovery for regulatory and legal compliance.

Purpose

The purpose of this research is to illustrate the complexities or impossibilities of e-discovery in cloud computing environments and determine if this knowledge can empower corporate customers to require discovery and forensic compliance by cloud vendors.

Significance of the Study

Cloud computing is an emerging paradigm that is experiencing widespread exposure and enormous marketing efforts in the information technology communities. Cloud vendors are working diligently to position themselves to take advantage of the potential market share of this technology. Literature in support of this paradigm promotes the potential for substantial savings in terms of networks and personnel, increased security, scalability, and even reduction of an organizations carbon footprint. What are not discussed in detail are the limitations or impossibility of e-discovery and digital forensic acquisition and investigation of enterprise data in the cloud environment. This paper will include literature reviews of expert commentary on the dangers of cloud computing with regard to e-discovery, and the need for verifiable proof that a cloud-hosting vendor will provide access to the network data required for digital forensics and compliance with e-discovery requests.

Literature Review

E-Discovery

Attorneys have performed searches for relevant documents for generations, combing through paper records in file cabinets, ledgers, logs and bank boxes. In many instances, the team producing the documents has "dumped" this information on the opposing party in an attempt to shift the burden of searching through voluminous records. Prior to the widespread use of computer systems, discovery of paper documents was typically possible to complete in the required time and considered an effective and virtually perfect way of producing information for litigation. Searching has always been a monumental task although modern computer systems have compounded these difficulties in staggering ways. Organizations create and store tens of thousands of times as many records as they did just two decades ago and this information is often stored in many places and varying formats. Pass (2011) notes that 487 billion gigabytes of data were created globally in 2008 and that this astounding number will double at least every 18 months. The practice of E-Discovery encompasses not only searching for relevant information using a wide variety of electronic search methods, but also the automation of these processes, manual review of electronically discovered data, as well as accessibility and retention policies and practices (Pass, 2011).

Electronically stored information (ESI) is information that is created and stored in digital form and accessible using computer hardware and software. ESI normally includes metadata, which is data about data. When documents, images and many other types of electronic data are created, information is added either to the digital file or in an attached/associated file such as the name, type, size of the file, location and ownership. Digital cameras often record the date and time of image creation and word processing documents contain metadata that often includes the

author, time and date of creation and the last date modified. Digital audio files contain the album name, song title, date recorded, and artist. Metadata is created and updated automatically by the computer application and may be created by the end user. E-Discovery experts must locate documents and files relevant to their cases, but also the associated metadata that contains valuable evidence about each file (Schuler, Peterson & Vincze, 2009)

The practice of digital forensics is also an important aspect of the E-Discovery process, since all information produced for litigation must also be defensible. Defensibility includes the people, search methodologies, collection methods, documentation, and chain of custody involved in the production of electronic data. Opposing counsel may question the methods of searching, producing, and chain of custody of files and documents. E-Discovery experts must therefore adhere to the sound and scientific practices of digital forensics in order to maintain admissibility of information in a court of law (Schuler, et al., 2009).

The Federal Rules of Civil Procedure (FRCP) that apply to all cases filed in federal court were amended on December 1, 2006 and included details regarding how electronically stored information (ESI) must be handled. Two notorious court cases, *Zubulake v. UBS Warburg* and *Coleman v. Morgan Stanley*, are considered catalysts for the FRCP amendment. Judges in these cases determined that data "spoliation will not be tolerated and that e-discovery responsiveness and due diligence are requirements" (Schuler, et al., 2009, Kindle location 456). The FRCP amendments included specific guidelines for the early attention and agreement on the types and formats for production of electronically stored data (ESI) as well as the systemized processes required to locate data. Parties are tasked to agree on the search keywords that will produce relevant information for the case, as well as other details including the common format for delivered data (i.e. paper files, .pdf, digital files, or forensic images). Pretrial and discovery

planning conferences are also requirements for opposing parties as well as strict search and preservation rules once litigation has begun. The FRCP amendments also include the word reasonable with regard to the efforts required for searching. This somewhat ambiguous term is the cause of much discussion in courts and the legal community. It also highlights the need to examine organizational data retention and destruction policies, since it is reasonable to search only data that exists as part of a regular retention policy. Organizations that historically retained backup tapes from old systems no longer in use for many years may reconsider these policies with regard to potential E-Discovery requirements. If an organization only retains backup information for a specific period, it is not reasonable to expect them to search and produce data outside of those retention parameters (Schuler, et al., 2009).

The Coleman v. Morgan Stanley case in 2004 is an example where sound retention policies may have been critically important. A Morgan Stanley IT executive testified that all relevant e-mail records had been provided to the court, when in fact 1600 unsearched backup tapes were discovered that might have contained important information. The court responded by awarding almost $1.8 billion dollars in damages to Coleman and determined that Morgan Stanley had committed fraud. The omission of the old backup tapes may have been an honest oversight or an attempt to conceal digital evidence, but it certainly illustrates the importance of proactively managing data and creating logical retention and destruction policies (Schuler, et al., 2009).

E-Discovery in the Legal and Technical Communities

E-Discovery combines the efforts of the legal, technical, and many other departments in a typical organization. There are significant costs and penalties if litigants violate the Federal Rules of Civil Procedures and organizations are becoming keenly aware of the importance of implementing effective discovery processes. It is important for organizations to be prepared to

respond to e-discovery requests in a manner that is accurate, reliable, cost-effective, and timely. Technology is now an important part of the E-Discovery process and can be used to reduce the time and effort required to provide documentation for review by the producing organization's legal counsel. A wide range of software systems are commercially available that not only facilitate the search processes, but also provide integrated data retention methodologies and forensically sound capture and segregation features designed specifically for E-Discovery (Schuler et al., 2009).

Organizations create, receive, and store information in a wide range of formats and locations. Data relevant to an E-Discovery case may include paper records, E-mail, digital telephone, audio and video files, engineering CAD drawings, social media information, and information from database systems. Information may also be stored in countless places and on many types of digital media. Many organizations own and manage their own data centers and regularly backup their data to removable media or offsite storage systems. Relevant data may also be stored in cloud environments, vendor-hosted environments, or on servers owned by social media organizations. Staff in various offices may have company data stored on their desktop computers, personally owned computers, removable media such as flash drives and CD's and even on their own Smartphone's. IT staff must maintain a clear and detailed map of where all of their data is stored in the event that these widely distributed repositories must be searched for litigation.

Organizations that create record retention policies must consider and map all possible data repositories since the scheduled destruction of information can provide enormous litigation cost savings if properly managed. Companies should avoid the necessity to explain to a court why information outside the retention period exists in some instances, and not others. Some

organizations are subject to specific retention periods for legal, contractual or other important reasons that must be factored into the overall plan. Mapping data locations and identification of data custodians can be an enormous task, however, if an e-discovery request is received, it is far better to have a documented plan and formal processes in place. The legal, technical, human resource and other departments may form a discovery response team to handle these tasks, with specialists who are responsible for various parts of the process. Litigants who can demonstrate structured policies with clearly defined retention and destruction periods are able to defend these policies in a court of law. If opposing litigants request information that is outside of the normal retention period, the organization cannot reasonably produce this information since it no longer exists (Schuler et al., 2009).

The combination of efforts between legal and information technology departments is necessary to determine the E-Discovery readiness for the organization. The legal staff understand the litigation requirements, and the information technology staff know where and how information is stored. Working together these groups can determine the systems that are available or must be implemented to support discovery, and ways that these systems can be leveraged to reduce costs, risk and speed of data production for litigation. It may also be important to investigate available external resources that may help to facilitate and respond defensibly to an E-Discovery request.

E-Discovery Searching and Web 2.0

The Federal Rules of Civil Procedure (FRCP) include specific rules regarding pre-trial activities. Rule 16(b) outlines the necessity for a Pretrial Conference where litigants discuss the disclosure provisions and any claims either party may have regarding privileged relevant information. Rule 26(f) of the FRCP mandates discussions regarding information preservation,

discovery of information and a mutually acceptable form of data production. Some litigants require paper copies of all information including relevant metadata for each item. Others may request .pdf formats, or even forensic images of the information. These early conferences and discussions also include the establishment of a list of search terms and date ranges that will be used in the search and discovery processes to locate relevant information.

"Data search and identification is the most important phase of the e-discovery process" according to Schuler, Peterson, & Vincze (2009, Kindle location 3877) since the requirement to protect, preserve and present relevant and forensically sound data in a court of law may be necessary for any organization. The demonstration that a litigant has made a "good-faith" or reasonable effort to locate and preserve data required for review is a somewhat gray area of the law but Schuler et al. (2009) note that "the courts have been quick to act when a party does not meet its preservation obligation" (Kindle location 3983). Litigants must provide relevant data including associated metadata in order to present the complete documentation requested. Metadata is both file and system data, which in the example of an e-mail would be the header information that contains the sender and recipient addresses, the IP address and other identifications pertinent to the transmission, plus the server logs that show receipt of the email transmission (Schuler et al., 2009).

Some organizations manage sophisticated fully integrated content and e-discovery management systems used to identify, preserve and manage data and documents used in litigation. These software systems provide advanced tools that allow legal and information technology professional's access to comprehensive search tools with the ability to mark discovered documents for retention. The search capabilities of some of the e-discovery software market leaders are capable of examining data from social networking sites, company database

and CRM applications and also electronic backup archives of email and other documents (Schuler et al., 2009).

James (2011) documents a webcast where E-Discovery professionals including Craig Ball, attorney and forensic technologist, discuss data mapping as the starting point for effective searching during the e-discovery process. According to James (2011) data mapping is the creation of a data inventory that is a critical element in the process, and is a guide of "where to look" for information that may be relevant to a case including paper, electronic data and documents, and voice mail messages. This process should involve checklists, in-person interviews with key custodians and always remembering to consider paper documents, home computers, and other places where corporate data may reside. Data custodians may forget data and documents that they produce in the course of their work so James (2011) notes that Craig highly recommends asking open ended questions and looking around the physical environment (James, 2011).

Searching for information often includes a number of different approaches including basic search techniques of custodian self-collection, and hard drive imaging. Electronic searches include keyword, metadata and Boolean searches. More advanced approaches include analytic technology that allows clustering and categorization, as well as concept and contextual searches. Organizations may have relevant data in their legacy relational database systems, structured and non-structured data warehouses, Web 2.0 sites such as blogs and social media websites and on backup tapes and offsite data repositories used for disaster recovery. Many data warehouse systems allow sophisticated searching for data based on concepts and contextual information and can be very useful for locating information. An unstructured data analytical engine is capable of locating abstract references, where a search for BAS might locate documents that the acronym

BAS is related to including ‘Basic Allowance for Subsistence', ‘Broadband Access Server’, ‘Basic programming language’ and/or ‘Bulgarian Academy of Sciences’, or things such as references to botany in fifteen languages (Inmon et al., 2011).

Schuler et al. (2009) caution readers that keyword, date and custodian searching may not be sufficient to provide and preserve relevant ESI since the many different search methods can produce "wildly different results" (Kindle location 4133). Each software application used in an E-Discovery case should be tested thoroughly for accuracy. Some software systems handle things such as abbreviations and polysemy more effectively than others do. Natural language synonymy also presents problems as do variations of words and names such as "Bob" and "Robert" and the possibility of foreign languages. Schuler et al. (2009) also comment that traditional keyword search engines are unable to discover data on alternate data sources such as video and voice recordings and may result in millions of irrelevant matches (Schuler et al., 2009).

Akers, Mason and Mansmann (2011) discuss the concept of analytic searching called "Latent Semantic Analysis" (p. 41). This technique includes production of documents based on both their content and keyword matches and electronically resembles the methods of human deduction from natural language, and which provides a weighting system for the results. The ideal E-Discovery search system according to Akers et al. (2011) would include efficient search capabilities able to analyze the full content of digital files and their metadata with comprehensive capabilities for weighting, analysis, validation and extraction of data. A virtual index stores the content and metadata separately and can grow exponentially while still allowing autonomous searches on either metadata or content. Sophisticated systems must be able to perform virtual

indexing to handle the large data sets in use today that may be many terabytes in size (Akers et al., 2011).

Although some sophisticated software applications are capable of searching Web 2.0 content, there are difficulties producing the associated metadata from these websites. Web 2.0 interactive collaborative websites such as blogs, social media websites and wiki's will not or cannot provide the complete metadata. However, this information is still being used in American courts, typically in printed or web archive formats, yet does not contain the actual associated metadata that is normally required for defensible data production (Meyer, 2009).

There are many different approaches to data identification and searching for relevant data during an e-discovery event. Some organizations employ outside legal and technical resources to assist with what can be an arduous process; others maintain a comprehensive mix of strict data policies and both content management and e-discovery software that is continuously updated and maintained. Regardless of a formal structured policy-based approach or one that is specific to a particular discovery case, the processes must be structured, well documented, and detailed enough to ensure that opposing council cannot claim there is a lack of participation and timely production of defensible data. E-discovery teams may place holds on many different types of data including forensic images, physical documents, and data files containing financial figures, e-mail documents, voice messages and even content from company social networking or web sites. The manual approach that some organizations take may seem cumbersome or even prehistoric to others who spend millions of dollars annually preparing for the eventual litigation request, but although the methods may be vastly different, the result of producing defensible, repeatable, and timely data is a necessity (Schuler et al., 2009).

E-Discovery and Social Media

The volume of data that is created on social media websites has reached astronomical proportions and is expected to increase substantially in the coming years. Frazier (2011) reports that Facebook users post over 30 billion individual pieces of digital content each month including comments, links, images, and videos. In 2010, over 25 billion tweets were posted on the popular Twitter website and that same year over 6 trillion text messages were exchanged. The lines between personal and organizational use of social and mobile media are typically very clear; however, many company-owned computers and mobile devices are being used by employees for non-company activities (Frazier, 2011).

Some organizations consider regulating employee use of social media yet this does little to stop their staff from engaging in social media discussions and posting updates to these popular websites through the use of their own Smartphone's and computers. Frazier (2011) comments that social and mobile media has been over-dramatized to some degree with regard to e-discovery and that "E-Discovery regulations do not pose an immediate and large threat to large corporations" (p. 1). Frazier (2011) notes that courts and E-Discovery experts distinguish between corporate and individual social media accounts allowing ample time for legal teams to consider a proactive and pragmatic approach to collecting this information.

Jaeger (2011, August) disagrees and provides statistics from an enterprise strategy group that 58 percent of organizational respondents already include social media applications in their E-Discovery collections including data from Facebook, Twitter and LinkedIn. Jaeger (2011, August) also reports that this same study reveals that cloud-based data "will parallel the expected rise in social media discovery" (p. 1).

E-Discovery preservation laws are very specific with regard to company owned data. Information must be produced for litigation regardless of where it resides, and companies must gain a better understanding of employee use of social media and extraction strategies to prepare for possible litigation. Jaeger (2011, August) comments that organizations must develop strategies for understanding, coexisting with, and managing social media data. The practice of E-Discovery is complicated by the use of social networking and mobile media in that this unstructured information is located created and stored in virtual, cloud-like, remote environments. Jaeger (2011, August) comments, "Social networking can exist anywhere making social media e-discovery a practical impossibility" (p. 1).

Courts have ruled in numerous instances that the data contained on social networking websites is discoverable. Social networking data can however be changed at any time and forensically sound metadata is not available. Legal counsel may rightfully then question the authenticity of this data raising unique E-Discovery concerns. Jaeger (2011, August) notes that the "archaic practice of taking screen-shots to capture images is no longer reliable" (p. 1) which challenges corporate legal departments to develop methods for real-time data collection.

The challenges of collecting data from social networking and other unstructured content sites such as blogs and wikis present business opportunities for software vendors to develop products to capture, monitor and collect this data. Jaeger (2011, August) comments that numerous software vendors and the social networking sites themselves are responding to these corporate challenges by offering solutions. Software-as-a-service vendors including Symantec sell products that allow archival of social media exchanges in electronic formats that can be used for E-Discovery. Jaeger (2011, August) quotes a senior executive from Actiance, a software vendor that provides these products who states, "Recording the content as people share it, and

retaining your own record of that data really plays to the regulators and the legislators, because you can't rely on Twitter or Facebook having that content available" (p. 1).

In the fall of 2011, Facebook released a new feature called "Download Your Information" that allows subscribers to extract a complete snapshot of their data, yet this small improvement does not included deleted data that might be critical for an E-Discovery investigation. Facebook subscribers can change their data at any time and the company does not permit forensic analysis or capture of historical or deleted member data.

Frazier (2011) provides details from a 2011 Compliance, Governance and Oversight Council (CGOC) survey stating that preservation and collection from mobile devices is typically performed only on company-issued devices. In these instances, only the server information is targeted and the phones and devices themselves are bypassed. This does not seem to be an extraordinarily comprehensive viewpoint, since many employees use their personally owned mobile devices for business purposes or to comment on business activities (Frasier, 2011).

An officially presided mock trial sponsored by the Compliance, Governance and Oversight Council (CGOC) in 2011 that included an interesting case where research and development employees had privileged information about a new drug they felt would not pass clinical trials. The employees posted information on social networking sites to this effect, and the company stock subsequently plummeted. The plaintiff requested the handheld devices, personally owned computers and social media postings of these employees for preservation and availability for examination. The costs involved in this investigation, just to locate information on these devices and websites, amounted to over $250,000. This example refers to an actual mock trial with an actual presiding judge. A judge in the Miami-Dade Circuit courts declared a mistrial with substantial awards including attorney's costs to the defense based on the plaintiff's

boss sending text messages to his employee while on the witness stand. The text messages were read to the court and included in the records of the trial.

Social networking websites provide an extensive resource for information gathering, and many law firms utilize this wealth of information to their advantage. The plaintiffs and defendants are not the only social network users who create a wealth of information for E-Discovery purposes, the attorney's themselves often have inappropriate behaviors that are used against them. Nelson, Simek, and Foltin (2009) list a number of attorney courtroom and litigation antics including making drunken unethical online defamatory comments and deceitful "friending" to gain access to opposing parties Facebook content. The use of discoverable content from social networking websites and mobile computing is expanding and may not be entirely defensible; however, it is currently being admitted into courtrooms across the nation even in forensically unsound formats.

Another popular online medium is Skype. This application allows international audio and video conversations and chat-based exchange of information using computers and smart phones. Over 20 million users may be active on this service during peak times with a staggering amount of digital information exchanged during these sessions. Skype supports file transfers, telephone audio calls, instant messaging, SMS and video chat conversations and records all of these activities in an unencrypted format. The instant messaging content, call logs, details of any file transfers performed during the sessions and complete metadata are stored on the local device. Message history is easily accessible even after exchanges are disconnected (Stuart, 2011).

One important aspect of the Skype service for E-Discovery is that all of the data is recorded in a SQLite database format that is fully accessible and can be indexed for straightforward data retrieval. These database tables not only store the instant message content

but also the metadata and complete details of file transfers allowing investigators to extract detailed evidence. Unlike many of the popular social networking websites, Skype provides all of the details in a non-proprietary format with great transparency. Information technology professionals must be aware that this information is stored on the local devices and consider ways to capture and integrate it in the regular backup and retention methodologies and policies if Skype is used in the workplace (Stuart, 2011).

Social network data does not fall into the commonly accepted definition of a cloud-computing environment, yet it is loosely considered to be part of "the cloud" since it is remotely hosted. This may be advantageous to cloud computing vendors who also do not or will not allow forensic practices in their multi-tenancy environments. This raises questions of why existing cases have allowed social network content in a court of law that in some instances is not accompanied by the important metadata that proves its authenticity (Stuart, 2011).

Challenges in E-Discovery

E-Discovery experts face many challenges while searching for, analyzing, and producing information during litigation. Even in a corporate-owned computing environment where data is reasonably accessible experts face difficulties with the enormous volumes of data. Traditional keyword search functionality on large data sets can produce millions of irrelevant "hits" necessitating more efficient and powerful search methodologies to handle the increasing volumes of data. There are also hundreds of possible electronic formats including paper documents making search and production of data with associated metadata a tedious and lengthy process. Data may reside in many places including employee handheld devices, removable media and personally owned computers as well as cloud environments, vendor-hosted servers, and international locations. Retention and legal hold policies are subject to scrutiny by the courts and

custodians can sometimes inadvertently destroy or compromise information required for litigation. Information technology professionals may also alter data by simply rebuilding indices and performing normal network maintenance that affects their files and data.

Jaeger (2011, December) notes that the greatest challenge for E-Discovery experts in 2011 is social media and cloud computing. Jaeger (2011, December) comments, "2011 will be a transformative year for electronic discovery, as social media becomes more common and cloud computing raises new concerns over the control and security of data" (p. 1).

Cloud Computing

Cloud computing is a "relatively new concept that offers the potential to deliver scalable electronic services to many" according to Reilly, Wren and Berry (2011, p. 1). Vogel (2011) disagrees with some amusement, remembering that this technological concept has been in use for many decades when remote connectivity to mainframe computers was in place in the 1960's. Vogel (2011) comments that "cloud computing is merely the newest label for the 1964 remote computing service originally called time-sharing at Dartmouth college" (p. 1), when dumb terminals connected to remote servers over telephone lines. Many organizations have offered hosted solutions for their products for years. The cloud computing paradigm is however seen by many as a new and emerging technology, and with good reason, since cloud computing is not just a remote hosted environment reminiscent of the early Dartmouth timeshares or software as a service (SaaS) remote application services that have existed for many years. It has distinct characteristics, service models and deployment models inherent to the term itself. There are many acronyms used in cloud computing to describe both the different types of architectures and available services (Reilly et al., 2011; Vogel, 2011).

Types of Cloud Environments

Regardless of the type of cloud environment, the United States National Institute of Standards and Technology (NIST) has created a standard definition of the elements that constitute a cloud computing environment that is mutually agreed upon by many cloud architects, designers and service providers. Mell and Grance (2011) describe the "five essential characteristics, three service models and four deployment models" in the NIST definition of cloud computing (p. 6). The essential characteristics include, "On-demand self-service, broad network access, resource pooling, rapid elasticity, and measured service" (Mell & Grance, 2011, p. 6).

On-demand self-service allows customers to manage their computing capabilities without any human intervention. Broad network access refers to the ability to access the cloud infrastructure using computers, workstations mobile computing devices and other standard browser-based devices. Resource pooling refers to multi-tenancy computing models where the computing resources are assigned dynamically based on customer demand. Rapid elasticity is the ability to quickly and automatically acquire or relinquish additional resources on demand, providing the customer impression of unlimited capabilities. Measured service is the automated control and optimization of computing resources providing a level of transparency for both the provider and cloud consumer (Mell & Grance, 2011).

The types of service models in cloud computing typically fall into one of three categories that include Software as a Service (SaaS), Platform as a Service (PaaS) or Infrastructure as a Service (IaaS). Software as a service has been available for many years and is a model where the provider owns and hosts software applications in a cloud infrastructure. With the exception of the ability to configure the application, the hosting vendor manages all of the application details.

The customer may be able to create new application users, add them to application security groups, and set some application configurations such as how dates are displayed, or selecting from cash versus accrual accounting methodologies. Platform as a Service (PaaS) represents a model where the cloud vendor provides the underlying infrastructure, and the customer deploys their choice of applications into the hosted environment and controls their systems and may also have some control over the configuration settings in their hosted environment (Mell & Grance, 2011). Infrastructure as a Service (IaaS) provides the customers with control over their networking environments and the ability to load and configure operating systems, storage, software and applications and may have limited control of networking devices and components such as firewalls and routers (Mell & Grance, 2011).

The NIST deployment models include Private, Community, Public, and Hybrid clouds. Private clouds are used exclusively by a single organization and their affiliates, and may exist within or remotely from the organization's premises. A community cloud is used by a specific community of consumers and may be owned and operated by an organization or a third party and physically exist in a variety of locations. The popular Apple iCloud is an example of a community cloud service, where consumers can store and manage their photos, information, applications, music and other Apple products. A public cloud according to Mell and Grance (2011) is available for use by the public and may be owned by a variety of organizations including academic institutions, government organizations or businesses. Hybrid clouds are a combination of two of the other three cloud infrastructures that individually retain their autonomy yet are integrated or bound by standard or proprietary technologies that enable sharing and portability. Mell and Grance (2011) provide an example of the hybrid deployment model where organizations may use a technique called "cloud bursting", that allows expansion into a

cloud environment for load balancing when the need for additional space or resources arises. An organization might have a private cloud and occasionally use the services of a public cloud to facilitate necessary on-demand expansion capabilities, by "bursting" their computing needs into a public cloud (Mell & Grance, 2011, p. 6).

A pubic cloud is a vendor-hosted shared computing environment where customers typically do not have access to the location. The provider is typically responsible for 100% of the capital expenditure to build and maintain the software and hardware that comprises the network. The customers normally share the operational expenses within a multi-tenancy environment. This paradigm allows customers to enjoy a scalable, networked computing environment without the overhead of maintaining it themselves. Customers can add new users and disk space quickly as needs and requirements change, and these types of cloud infrastructures typically provide outstanding support for mobile computing devices since a web browser is the only interface required to access the applications. An example of a public cloud is Amazon's AWS services (Swindler, 2011).

A private cloud is a customer-centric infrastructure that is completely dedicated to a single organization. It is the equivalent of in-house development of or procurement of a vendor provided remote computing infrastructure if it contains the characteristics outlined by Mell and Grance (2011, p. 6) that include on-demand self-service, broad network access, resource pooling, rapid elasticity, and measured service. An organization may state that they are running a private cloud if they are using the technology of virtualization. Virtualization is the creation of an abstraction layer, effectively 'fooling' the operating system into believing there is a logical view of computing resources that might span multiple physical servers. The earliest forms of virtualization were on mainframes where physical servers were divided so that they could run

multiple operating systems. This was useful for testing new applications on differing platforms where developers could see how a new application would run in another environment. This technique allows the consolidation of servers to reduce the overhead of computing infrastructures and allows IT managers to consider the services that existing technology can provide. If this virtualized environment contains automated and managed resource provisioning as well as an automated and integrated billing or chargeback system, then this environment is maturing into a private cloud (Mell & Grance, 2011).

Emerging Challenges from Cloud Computing

The marketing of cloud computing is spreading quickly to consumers and businesses and 'the cloud' is becoming a common household term. Non-technical Apple Smartphone and computer users ranging in age from children to adults are routinely offloading data to their iCloud space. Information technology executives are discussing cloud technologies in the boardroom and break room as this term becomes ubiquitous. A fascinating example of visual semiotics occurred when observing a pre-school child who noticed a cloud sticker in an elevator and excitedly announced, "Look! iCloud!". Pelz-Sharpe (2011, p. 1) comments, "Cloud computing as a term, has been overhyped to a degree not seen since the nonsense of the dot-com era, yet beneath the hype there is value to be found".

Pelz-Sharpe (2011) continues by commenting on the widely held theory that cloud computing is always substantially more cost effective than traditional network infrastructures, yet has witnessed several instances where organizational analyses have resulted in the cloud as a more expensive option. Other considerations include the possibility of failures. In-house infrastructures typically have a single failure point that might be vulnerable. The cloud "removes the very concept of a single point of failure, replacing it with multiple potential failure points"

(Pelz-Sharpe, 2011, p. 1), yet what happens if an entire data center is compromised containing replicated data? Pelz-Sharpe (2011) aptly notes that cloud computing has a significant future in the world of enterprise content management and computing, but has a "nascent set of services that are not particularly well-defined or understood by either the service providers or those seeking these services" (Pelz-Sharpe, 2011, p. 1). The marketing hype promoting the many advantages of cloud computing is both distorting and obscuring the available options and patently ignoring the risks (Pelz-Sharpe, 2011).

Risks in Cloud Computing

Security is a great concern for organizations considering a public or private cloud environment. Golden (2011) notes a survey from April 2010 where "45 percent of respondents felt the risks of cloud computing outweigh its benefits" (Golden, 2011, p. 1). Hickey (2011) writes about a survey by Norwich University where the vast majority of 646 respondents from federal, state, and local government and higher education organizations are hesitant to employ public cloud models with only 6.9 percent supporting its viability. More than sixty-eight percent of these respondents are in favor of private or hybrid cloud environments and many state that security issues, costs of moving existing infrastructures to the cloud and exit strategies from cloud providers are barriers to adoption of this technology (Golden, 2011; Hickey, 2011).

Golden (2011) attempts to dispel the concept that public clouds are less secure than private ones and states that this misunderstanding is "an overly simplistic characterization", and "arises from the fact that cloud computing contains two key differences from traditional computing: virtualization and dynamism" (Golden, 2011, p. 1). . Golden (2011) explains a fundamental cloud computing concept of a hypervisor that insulates applications from traditional security tools that inspect network traffic for security threats. Virtual machines on the same

servers can send valid as well as malicious information through the hypervisor to communicate with one another and effectively bypass the physical network and its security mechanisms. A private cloud does not prevent or protect against these security threats (Golden, 2011).

Cloud computing according to Golden (2011) requires an automated environment in order to achieve the "agility and elasticity to respond to changing application conditions by moving virtual machines quickly and spinning up additional virtual machines to manage changing load patterns" (Golden, 2011, p. 2). Cloud environments must therefore automate the installation and configuration of software and security without any human intervention. Golden (2011) comments that in a private cloud, automation and alignment of processes may not be present, which presents a substantial vulnerability.

Another common misperception about public clouds according to Golden (2011) is that the cloud service provider (CSP) is entirely responsible for sound security practices. This fortifies the prevailing theory that a private cloud is more secure. The CSP is responsible for the security mechanisms and methodologies from infrastructure to the point of interfacing with the applications and hosting environments. The public cloud customer shares the responsibility for security for their users interfacing with the cloud environment and their individual applications. The security at the application-level is a point of vulnerability that no CSP can be expected to have responsibility over. Golden (2011) reinforces that there is a dynamic partnership between a client and a CSP in securing the overall computing experience and considers it reckless and irresponsible for any organization to assume that all security lies with the CSP. Golden (2011) also comments that if any type of cloud application is poorly configured it can be vulnerable to security risks and that "charactering security risks in either situation as black or white is simplistic" (Golden, 2011, p. 4).

Gartner researchers Heiser and Nicholett (2008) address recommendations for assessing the potential risks of adopting a cloud computing methodology for an organization. The assessments necessary to analyze security, privacy and regulatory compliance risks are compounded by the fact that a cloud provider might actually be part of a group operating as a single service and not have ultimate control over the technology, physical locations, or staff involved. In this scenario, each of the subcontractors involved might have access to unencrypted data on its servers making it exponentially more challenging to assess the inherent risks. Heiser and Nicholett (2008) detail eight specific areas to evaluate when considering a cloud service provider including, "privileged user access, compliance, data location, data segregation, availability, recovery, investigative support, and viability" (Heiser & Nicholett, 2008, p. 1) . An additional area not specifically mentioned is the technological considerations for a termination agreement (Heiser & Nicholett, 2008).

Privileged User Access and Compliance

The cloud provider employees may have direct access to data stored on their systems, which is the privileged user access risk. Heiser and Nicholett (2008, p. 2) note that there is an "inverse relationship between loyalty and risk" and stress that individuals without a long-term commitment or involvement with an organization pose a higher risk. In-house employees would naturally garner a higher level of trust than those working for a cloud provider. Therefore, organizations should inquire about the hiring and supervision of cloud employees and subcontractors as well as the specific controls and monitoring for their access. Compliance is another security risk that concerns many organizations. Publicly held organizations must comply with the requirements set forth in the Sarbanes-Oxley act, health data must comply with HIPPA privacy requirements and many other regulatory considerations exist in the area of compliance.

Cloud providers under consideration must also contractually be willing to submit to external audits and security certification standards (Heiser & Nicholett, 2008).

Data Location and Data Segregation

The physical location of the data is a concern for many organizations since national and other regulatory bodies may prevent the storage of personal information outside of a specific jurisdiction. These concerns are especially important in the European Economic Area (EEA, that includes the European Union plus Iceland, Liechtenstein and Norway) where according to Hon, Hornle, and Millard (2011) the European Union Data Protection Directive prevents EEA organizations from storing personal data on non-EEA cloud providers systems. The physical location of the cloud providers is of critical importance to organizations in these countries since the 'Directive' is very specific about what are considered adequate levels of protection for their citizen's privacy.

Heiser and Nicholett (2008) also elaborate on the considerations of data segregation in the cloud, and note that most providers use Secure Socket Layers (SSL) technologies to protect data during transmissions. They however advise customers to inquire on the procedures to segregate data when in storage, and note, "if your data can be read at your provider's site, then you have to assume that it will be read" (Heiser & Nicholett, 2008, p. 3). Many providers advertise encryption methodologies, which is admirable and normally quite safe. However if these implementations were not designed and tested by certified professionals performing structured protocol analyses, encryption malfunctions can render data as completely unusable and complicate availability. This is especially important with regard to e-discovery and digital forensic investigations. Further, a cloud vendor should be forthcoming with a list of the staff that

have access to the decryption keys and methods of contacting them in the event of an encryption emergency (Heiser & Nicholett, 2008).

Availability and Recovery Considerations

Availability is a widely marketed benefit of cloud computing, although Oreskovic and Wohl (2011) note that disruptions to the Amazon EC2 data center hosts caused significant disruption to social networking sites including Foursquare and Quora who reported failures and latency issues due to capacity issues with the Amazon services. Reddit, an organization that reports over a billion monthly page views on their Amazon-hosted website is considering the costly move from Amazon Web Services cloud hosting to their local systems to improve reliability according to Clark (2011). Heiser and Nicholett (2008) state that guaranteed availability is not typically provided by the cloud-based vendors, which would threaten critical business applications for their customers. Heiser and Nicholett (2008) recommend establishing contractual penalties when levels of service that are not met. In-house information technology teams typically have standard procedures to mitigate networking issues and notifications in place to alert staff when problems arise. Organizations must achieve some assurances that cloud data center staff will have similar levels of commitment to rapid problem resolution (Heiser & Nicholett (2008).

Disaster recovery plans are important for any organization, and equally important for cloud providers. Cloud vendors may replicate data and computing infrastructures among multiple data centers to mitigate vulnerabilities from a total system failure. Heiser and Nicholett (2008) recommend obtaining complete details of the disaster recovery plans, even when cloud vendors are not able to or refuse to state where a customer's data is physically stored (Heiser & Nicholett, 2008).

Viability and Investigative Support in Cloud Computing

The viability of the cloud hosting organization is an obvious concern, but one that may be overlooked. Heiser and Nicholett (2008, 4.) recommend posing questions such as, "What happens if the provider goes broke, or if it does, how will you be able to use your data or get it back?" These are excellent questions, and ones that must be contractually specified. This author would consider asking what would happen if the data centers experienced a major catastrophe such as a tsunami or were located near an international crisis. Cloud providers are typically responsible for backup strategies and replication of data, but the specifications of the format and media used in these processes is of interest from a security perspective. Organizations must ask where backups are stored and the details of how they are secured. Organization may ask what would happen if they wished to have data off-loaded to another in-house data center upon termination of the cloud commitment. Heiser and Nicholett (2008) neglected to include termination agreements and strategies, but these are important additional questions and concerns.

Hickey (2011, p. 1) notes that more than ninety percent of federal respondents "don't have or aren't aware if they have a cloud computing exit strategy, meaning if they are uncertain if they can move their data or change cloud providers". Knowing that cloud providers replicate and move data continuously between the many available resources and that even deleted data can be forensically extracted from disk, exit strategies must consider data that may be left behind on backup media, or on widely dispersed computing platforms. An associated security question would include provisions for subcontracted cloud partners who terminate their agreements, and the procedures and assurances that customer data will be securely purged and permanently erased from their machines (Heiser & Nicholett; Hickey, 2011).

Investigative support is a critical concern since Schuler, Peterson and Vincze (2009) state that the need for e-discovery preparedness practices in response to a lawsuit is almost inevitable for any type of organization. Heiser and Nicholett (2008) elaborate on the necessity for organizations to perform internal investigations or e-discovery practices when there are suspicions of illegal or inappropriate activities or there are requirements to produce relevant data for litigation. These noted Gartner researchers stated, "If you are considering purchasing a service that would process anything considered a business record, or if you otherwise anticipate a need to conduct investigations, then you cannot assume that a service provider will be willing, or even able to support them" (Heiser & Nicholett, 2008, p.4). Virtually any type of data, including e-commerce transactions, website dialog with customers, e-mail and stored documents are considered business records. Heiser and Nicholett (2008) continue by adding that cloud services are particularly difficult to access forensically since important network logs and data are combined or collocated and may also be located across a range of changing data centers and servers. Digital forensics and e-discovery according to Heiser and Nicholett (2008) are arduous and costly pursuits when performed on company owned networks and systems that investigators can access. When organizational information is stored in a cloud environment, digital forensics and e-discovery becomes exponentially more difficult, or impossible. These authors conclude their narrative about investigative support with the statement, "If you cannot get a contractual commitment to support specific forms of forensic and e-discovery investigation, along with evidence that the provider has already successfully supported such activities, then your own safe assumption is that investigation and discovery requests will be impossible" (Heiser & Nicholett, 2008, p.4).

This single statement containing the finality of the word 'impossible' as stated by Heiser and Nicholett (2008) with regard to e-discovery in the cloud, is the reason that the investigation into this problem began. How then would any organization be able to justify the risk of cloud computing, if it might be impossible to investigate criminal activities or provide forensically sound data when faced with court-ordered e-discovery evidence production required for litigation? An attorney's response to court-ordered litigation requests cannot include, "Sorry, Your Honor, my client stores data in the cloud, and therefore we cannot produce the requested documents and data required for this case".

Digital Forensics in the Cloud and Applications to E-Discovery

Cybercrime is escalating annually and according to Weigel (2011), these crime rates have risen three percent with an estimated loss of revenue in the United States of $139 billion. The most frequent types of cybercrime are malware, online credit card fraud and phishing, yet insider attacks and corporate espionage are also on the rise. A terminated or disgruntled employee can easily leave their place of employment with valuable information. Computer Crime Research Center Staff (2011) reference a case of corporate espionage where a disgruntled employee walked out of a building with the plans for a new car model on a flash drive. When details of the new design were published, the organization reportedly lost an estimated billion dollars as sales of existing models halted when consumers elected to wait for a new model. Many of these statistics will never be revealed in cybercrime statistics since organizations often do not wish to publicize security events for a variety of reasons.

Digital forensics practices allow the formal reconstruction of electronic crimes when investigators gather evidence used to prove or answer the 'who, what, when, where and why's' of an event. This scientific process is performed either on static computers that have been powered

off, or on live systems or networks that remain live and functioning. Live digital forensics is becoming a common practice, since investigators can take advantage of valuable evidence located in computer memory, cache and other system locations that might be lost if the power is removed. Since live network forensics have become commonplace, one might wonder why a Cloud networked environment poses such extraordinary challenges. The practice of e-discovery requires the same preservation of evidence that digital forensic investigators strive to achieve, since evidence must be irrefutable and repeatable if possible. The problem with live forensic investigations of Cloud computer environments is that these systems are shared by a variety of clients. Cloud customers in a distributed, virtualized, multi-tenancy environment must be guaranteed privacy of their data and information systems, and data for multiple customers may be co-mingled on a single disk, or spread among many servers that are widely geographically dispersed. The virtual cloud environment provides a scalable, on-demand resource that must still be subjected to digital forensic investigations.

Garfinkel (2010) details many of the challenges for digital forensics experts and notes, "digital forensics is facing a crisis" (2010, p. 3) based on advances in the computer industry. The increasing size of storage devices requires an immense amount of time to create forensic images. Flash storage devices and embedded hardware memory poses challenges to find data and storage devices that cannot be removed or imaged. Encryption techniques often prevent analysis of recoverable data. Garfinkel (2010, p. 3) comments , "Use of the cloud for remote processing and storage, and to split a single data structure into elements, means that frequently data or code cannot even be found". Garfinkel (2010, p. 3) states, "Cloud computing in particular may make it impossible to perform basic forensic steps of data preservation and isolation on systems of forensic interest", further reinforcing similar comments by Heiser and Nicholett (2008).

Delport, Oliver and Köhn (2011) describe the need for digital forensics in cloud environments as well as the complications these environments bring to this practice. Delport et al. (2011) use the term "instance" in reference to a "virtual system resource established within a cloud" (p. 1) and note that a single node can contain multiple instances. An instance might be the backup of data for a single user, or a large e-commerce website with a database used by an organization. They elaborate by stating that the Cloud is comprised of multiple nodes with no predefined boundaries. Delport et al. (2011) stress the need to isolate a potential crime scene to protect and preserve evidence from contamination. Digital forensic and E-discovery practices require strict, methodical, and scientific procedures to gather and isolate information since opposing counsel will undoubtedly question the procedures used to gather evidence. If the continuity and chain of custody for any piece of data cannot be proven, the information may be rendered useless and inadmissible in a court of law.

Delport et al. (2011) detail various methods and techniques that digital forensic investigators might employ when attempting to investigate a cloud instance including "Instance Relocation, Server Farming, Failover, Address Relocation, Sandboxing, Man in the Middle (MITM) and Let's Hope for the Best (LHFTB)" (p. 3). All of these techniques have inherent risks and each organization must determine the acceptable levels of risk that may be present in each methodology.

Instance Relocation

Instance relocation is the process of moving the three distinct units that comprise an instance which are the data stored on secondary storage, the complete contents of the virtual memory, and the running processes. Instance relocation can be part of an automated process, or performed manually by an administrator or investigator.

One manual relocation method is to end one instance and create a new one. The storage is copied to an image file using standard forensic imaging tools such as EnCase or dd_rescue, and the virtual memory is also written to a file. The newly created instance might be created with the identical network address but contained on a different node. The risks when using this method include the running processes, which are difficult or impossible to store in a way that they can be restored or restarted since they are in a constant state of change. Another manual relocation method is to create a new instance and then move the various information units to the new instance and delete the old one. The running processes can be moved with this method using established, tested procedures, but both of these procedures contain the risk of losing virtual memory (Delport et al., 2011).

Automated instance relocation is commonly tested and proven by the creators of the cloud environment and can be part of the cloud operating system design. These automated processes are used for load balancing, conflicts between co-located instances, and administrators requests to move them for digital forensic investigations. Cloud providers should be required to maintain a customer's confidentiality, integrity and availability of their data.

The Automatic Instance Relocation techniques that are part of the cloud operating systems are thoroughly tested since this is how the cloud environment balances the daily loads on the resources. If an instance becomes extremely resource-intensive it may be automatically moved from one node to another and the prior node cleared. If these mechanisms are not reliable, the customers may experience service interruptions or data loss. Delport et al. (2011) note that even these automated processes carry risks since copying the storage media is very simple. However, keeping the instance running, especially when involved in transactional operations is difficult and risky. Digital forensic investigators might intentionally be able to

overload a particular instance, forcing the automated balancing techniques resulting in clearing a node for an investigation. Delport et al. (2011) also comment that the automated balancing may compromise the forensic investigator who may not be able to prove the reliability of their methods if the cloud operating system enacts the movement of the instance.

Regardless of the type of relocation technique, these instance relocation methods are extremely difficult to implement and fraught with risks including the violation of the data and transactions contained in the instance and integrity of a forensic investigation when trying to prove methods and chain of custody of the data movement. Both the manual and automatic movement of instances includes risks. "Storage media can easily be copied but it is a non-trivial task to send it to the new instance and keep that instance running" (Delport et al., 2011, p. 4) since the overwriting of system files and data can result in the failure of the new instance.

Server Farming

A server farm is a multi-node system where an application such as a website is split over multiple nodes to ensure quality of service for online customers and website visitors. If one node fails, the other remaining instances will function providing additional website quality of service. In order for an organization to perform server farming in a cloud environment, the cloud infrastructure must contain the ability to reroute network traffic. If this feature exists, digital forensic investigators have the ability to terminate a node that contains suspicious information or traffic and isolate it for investigation. Server farming is an expensive option in terms of both system overhead and creation for a provider, yet may offer cloud customers an outstanding capability to implement digital forensic investigations. However, "If the implementation is wrong, the digital forensic investigation can result in the loss of confidentiality, integrity and availability" of customer data (Delport et al., 2011, p. 4).

Failover

Failover is a technique used to provide high availability for websites and other mission critical applications where one or more servers replicate the production server creating a backup that can immediately take over if a primary server fails. Delport at al., (2011) list three types of failover techniques including client-based, DNS-based and IP-address take over systems. The client-based systems have at least one primary and backup server that the client is aware of and can access if system difficulties arise. The DNS-based failover is automatically directed to a backup by the DNS server if a primary server fails. The IP-address takeover method alerts a secondary server that takes over if a primary one fails. Digital forensic investigators could make great use of the IP-address failover methodology by terminating the primary instance to use for forensic investigations while the backup continues to provide a viable and reliable instance for continued operations. If the IP-address failover feature is available in the cloud environment, the digital forensics team must rely on the cloud operating system manufacturers for assistance and there is still a distinct possibility of data loss and compromise to the integrity and confidentiality of information (Delport et al., 2011).

Address Relocation

Address relocation requires two DNS servers running in the cloud and might be considered a waste of resources by the cloud provider or not available as an option in the infrastructure. When using this methodology, network traffic is rerouted without end-user knowledge, and appears to route from the original computer as part of a failover method. This feature may be part of the cloud service where a new instance and traffic are re-routed to this instance; however, the replication must be correctly implemented in order for this method to work without data loss. This technique may allow a digital forensic investigator to isolate the

original instance while the new one is processing traffic. It is risky and must be tested thoroughly in order to ensure its viability as a forensic option (Delport et al., 2011).

Sandboxing

Sandboxing is another forensic technique that can be used to isolate an instance in a cloud environment. It creates an isolation so the instance cannot interact with any resources or other instances outside of the sandbox environment. This technique is performed using either an application launched from the cloud operating system or by an investigator who launches the application from the same network on the instance itself. There is significant risks involved with this technique, since the instance may recognize that it has been 'sandboxed' and tamper with or destroy valuable evidence. It is also difficult to block the normal network traffic in a forensically sound manner resulting in inconclusive or spoiled evidence. It can however aid investigators in removing other instances from a cloud node while the target instance is in a sandboxed isolation (Delport et al., 2011).

Man in the Middle

If a cloud vendor permits, a man in the middle forensic investigation can be launched when an entity is placed between the cloud hardware and the instance so that data moving between the two can be analyzed. This technique can remain latent until there is suspicion of wrongdoing and when activated can prevent the suspect instance from deleting data and altering RAM. The instance does not recognize the monitoring of the middle entity and can function normally without interfering with other instances. The problems lie with the ability to implement this technique since it is reliant on the cloud operating system manufacturers who may not wish to implement reverse engineered software on their operating systems since there are significant legal liabilities. The process is also not forensically sound since reverse

engineering operating system software violates the software terms of use and there it is unlikely that the reverse engineered version can be proven in a court of law (Delport et al., 2011).

Let's Hope for the Best

The last technical forensic scenario presented by Delport et al. (2011) is the "Let's Hope for the Best" option, which requires the termination of the node and subsequent forensic imaging of the drives. This method again poses difficulties in that the drives have shared tenancy and a cloud provider must protect the information of other customers who have data on the same physical drive. This practice violates the confidentiality of the other tenants and is an unlikely possibility that the cloud provider would allow this practice. Since data in a cloud environment is replicated across multiple servers in often widely geographically dispersed locations, there may be dozens of additional tenants data contained in drives. Delport et al. (2011) note that even if a cloud vendor permitted imaging of the physical drives containing data belonging to other organizations, piecing together an original virtual hard drive would likely lose valuable evidence. The RAM and other valuable from a list environment would also be compromised or permanently lost (Delport et al., 2011).

All of the techniques presented by Delport et al. (2011) are ones that are technologically possible, yet contain significant risks and may be impossible due to the reliance on the cloud vendor or the cloud systems operating system manufacturer. Instance relocation is risky since running processes, virtual memory and overwriting of system files and data is likely, and the technique is reliant on the cloud vendor permitting imaging a drive that is likely shared with other customers. Server Farming has a significant overhead and it is unlikely that a cloud vendor would allow rerouting of network traffic. Failover may not be possible since it relies on the cloud vendor and operating system vendor and poses risks to data loss and potential compromise

to the integrity and confidentiality of customer's information. Address relocation is risky, requires two DNS servers, and may be a waste of resources for a cloud provider and therefore not available. Sandboxing is not forensically sound, and poses challenges when blocking network traffic that multiple customers are likely to use. Man in the Middle requires reverse engineering operating system programs and is likely to be allowed or supported by the cloud vendor since reverse engineered applications are not forensically sound. The 'Let's Hope for the Best' method relies on the cloud vendor permitting the imaging of an entire shared drive or piecing together information from a distributed virtual drive, losing RAM and running program information in the process. Since it is unlikely that a cloud vendor would authorize many of these techniques and that most of them contain significant risks, there do not seem to be many areas for forensic analysis in a cloud environment. E-discovery relies on sound digital forensic practices and therefore Delport et al. (2011) have further illustrated the in-depth technical reasons why forensics in the cloud is impossible, unlikely to be permitted, or may present challenges that would compromise any investigation.

Birk and Wegener (2011) add that digital forensic data in cloud infrastructures poses a "huge problem" since investigators never know where the data actually resides and there is no access to the physical hardware. Birk and Wegener (2011, p. 2) also state, "The snapshot technology presented provides a powerful tool to freeze system states and thus makes digital investigations, at least in infrastructure as a service (IaaS) scenarios theoretically possible". The words "theoretically possible" are not encouraging for those considering moving production systems or critical information to a cloud environment. Many who consider cloud computing as a cost-effective option are seeking the public cloud implementations that do not include the "theoretically possible" investigations noted by Birk and Wegener (2011). The amount of

technical evidence that an investigator might access is also "strongly divergent" between the various cloud server and deployment types according to Birk and Wegener (2011), who also note that digital forensics may need to be revised and adapted to the emerging cloud environments. This will certainly rely on the cloud providers to provide the tools and capabilities to accommodate forensic practices in their environments.

Birk and Wegener (2011) also state that access to metadata in any cloud environment is a great concern but in Software as a Service (SaaS) environments, it is a "huge problem" (p. 6). The cloud service providers do not offer any possibility for their customers to determine if data has been accessed or compromised, by a valid user or an adversary. There is therefore no ability to prove who accessed data, which is a critical element in any e-discovery case. It is not until the cloud service providers create a provenance detection mechanism for these distributed environments that this will be possible. "Cloud service providers should have the ability to receive specific information such as access, error and event logs that could improve their situation in case of an investigation" according to Birk and Wegener (2011, p. 7). Other specific suggestions offered on how cloud service providers might implement this cryptographic functionality include "Provable Data Possession (PDP), or Proofs of Retrievability (POR)" (Birk & Wegener, 2011, p. 7) yet no existing cloud service providers offer these features.

Cloud Technology and E-Discovery

E-Discovery relies upon and is based upon sound forensic practices. "The lack of knowledge about the internal processes, infrastructure and system components make the usage of current cloud computing offers a game of hazard" according to Birk and Wegener (2011, p. 7). The risks and implications of this statement would prevent any legal counsel from recommending or perhaps even permitting their organizations to consider a cloud environment.

There is a considerable lack of transparency, which is required in order to produce information for e-discovery purposes. Even in the "Infrastructure as a Service" cloud computing environments, an acceptable level of transparency is not present. Birk and Wegener (2011) consider this one of the primary reasons that cloud computing is not reaching its potential. In many instances, the determination of the body of law that governs or restricts the scope of an investigation can pose many challenges. What might be a crime worthy of investigation in one country or jurisdiction might not be a crime where the data is physically located in these distributed and highly replicated environments (Birk & Wegener, 2011).

A digital forensic investigation requires piecing together the timelines and actors in an event. In an e-discovery case, opposing counsel will undoubtedly wish to see a clear and forensically sound picture of events. Cloud service providers do not wish to make the network logs and information about the underlying infrastructure available since it could be used to compromise their systems by attackers or by competitors to better improve their services. E-discovery experts cannot provide chain of evidence or even a complete transactional picture of how data was created, altered or potentially compromised since logs and metadata are not available from the cloud providers (Birk & Wegener, 2011).

Cloud vendors might profit from the implementation of forensic capabilities since they are so obviously missing from the current offerings. Birk and Wegener (2011) make some suggestions of ways that these vendors might implement successful solutions. Network logs are often co-mingled in multi-tenancy cloud environments and a read-only application programming interface (API) that provides customer access to network, process and access logs might offer a level of trust in cloud computing as a choice for their data.

Another issue that e-discovery investigators may have with cloud computing is the assurances that information is deleted. Schuler et al. (2009) discuss in detail the necessity for a defensible retention policy for e-discovery purposes. Organizations that can prove they have a regular and documented retention policy are less likely to be sanctioned for misconduct with regard to production of information that was deleted according to this policy. The Federal Rules of Civil Procedure, specifically rule 379f called the "safe harbor" rule, state that courts may not impose sanctions for failure to produce electronic information that was lost as a result of routine business practices of data retention. If the cloud service provider does not effectively allow customers to adhere to this practice and information that was meant to be deleted according to policy is discovered, it can open up a myriad of issues in an e-discovery case. If opposing counsel wonders why some records exist outside of normal retention policies, they may question what other valuable relevant information might still be available. Birk and Wegener (2011) also wonder how cloud service providers can ensure their customers that email, and other potentially sensitive data on a virtual machine have been deleted with no remaining traces including metadata. If cloud computing vendors cannot provide verifiable levels of trust and data integrity, cannot or will not allow their customers to launch forensic investigations, and cannot provide the live or static forensic information required to perform e-discovery investigations on their systems, they remain non-viable computing options for organizations (Birk & Wegener, 2011; Schuler, et al., 2009).

When considering the following litigation statistics, the impossibility of many forensic and e-discovery practices in a cloud computing environment severely lessens an advisable impetuous to move to these environments until providers commit to assist with the challenges their customers are likely to encounter.

- Over one-third of corporations in the United States are facing at least twenty-five individual lawsuits.
- More than eighteen percent of U.S. companies are currently in litigation in more than 100 cases.
- Forty-eight percent of organizations reporting encountered new regulatory requirements and proceedings in the previous year.
- Forty percent of U.S. companies have encountered at least one instance of litigation that involved more than $20 million dollars (Schuler et al., 2009).

Hunter (2011) disagrees, and states that organizations can "easily manage the risks of altering metadata and risks of violating international privacy laws" (Hunter, 2011, p. 1) by requiring cloud providers to include mandatory statements in their agreements. These include provisions that none of the data may be stored outside of the United States, specifically addressing how metadata will be provided in the event of litigation and how it will be stored in the cloud environment. Hunter (2011) however also notes that cloud technology must be improved in order to facilitate these requests and that cloud administrators do not currently have methods in place to respond to e-discovery requests. Hunter (2011) also comments that if these features are present and available by cloud providers that e-discovery might offer savings over traditional methodologies. Statements including the word "if" with regard to cloud vendors replacing the "hodgepodge of hard drives, servers and removable storage that now house our data" (Hunter, 2011, p. 1), still force a conclusion that major changes are required before the cloud can emerge as a viable solution that supports e-discovery.

Methodology

Subjects and Setting

This major research paper did not include any subjects or take place in a particular setting.

Data Collection Technique

The inspiration for this paper was based on a single comment in the paper entitled, "Assessing the Security Risks of Cloud Computing" by Heiser and Nicolett (2008). They stated, "If you cannot get a contractual commitment to support specific forms of investigation, along with evidence that the vendor has already successfully supported such activities, then your only safe assumption is that investigation and discovery requests will be impossible" (p. 4). The word "impossible" in relation to computing of any kind is a very strong statement, and sparked an tremendous interest and extensive resultant research into cloud computing and a keen interest in the reasons that these Gartner researchers may have made this statement in 2008 (Heiser & Nicolett, 2008).

The majority of the literature included in this paper is less than a year old at the time of writing, and a number of other scholars still concur with Heiser and Nicolett's (2008) statement, especially for public cloud computing environments. The collection of literature was based on a chronological approach, reviewing dozens of scholarly papers and articles in technical and legal magazines and journals from 2008 to present. This approach was used to determine how and if cloud computing environments and vendors have changed during this time with regarding to allowing true forensic investigations.

Statistical Analysis

This study would lend itself to a detailed quantitative analysis of cloud providers and their responses to a detailed E-Discovery survey asking questions regarding privileged user access, regulatory compliance, data location and segregation, forensic capabilities, security and recovery. This would provide organizations considering a cloud computing environment with the necessary information to make informed choices to secure, access, and perform structured, scientific forensics and E-Discovery on their data.

A qualitative analysis of the same cloud providers would reveal the underlying contractual and operational possibilities that their potential customers could realize if they moved their computing requirements to their cloud platforms.

This combined study would support or deny the position that cloud computing does not support E-Discovery and therefore is currently an unsuitable choice for organizations who must consider their governance strategies.

Limitations of the Study

This paper has limitations including the lack of significant volumes of research into the difficulties of E-Discovery in cloud computing environments, and the time to perform a complete statistical analysis of current cloud provider offerings and capabilities to allow forensics and E-Discovery on their systems.

Discussion

Cloud computing is a new and developing field that may open many possibilities for cost savings and expandable services for organizations. It reduces or eliminates the need to create and maintain enterprise network systems and may provide a scalable, secure and flexible computing platform. The limitations in this study include the inability to perform a full statistical review of

the existing cloud providers and learn more about how organizations might perform digital forensics and E-Discovery on these systems. There is currently no statistical information to support the literature included in this paper. The cloud computing infrastructures are new and emerging and the widespread marketing of the benefits of these environments is far more prevalent than the risks that are indicated by the literature included in this paper.

Review of the Problem or Research Question

The apparent problem that E-Discovery is substantially more difficult in all of the cloud deployment models including Private, Public and Hybrid is confirmed based on the review of the literature. The statistical analysis of cloud provider offerings would be an ideal compliment to this study and would provide the confirmations required for this supposition. Cloud environments are remotely located with widely dispersed data and pose significant challenges in isolating instances in order to perform the scientific digital forensic techniques required for a defensible E-Discovery project. The literature review has confirmed the initial problem and several technical scholars concur that E-Discovery should be considered impossible until such time that cloud providers contractually ensure their customers that E-Discovery is supported.

Summary of Literature Review

Cloud computing is an exciting, emerging computing possibility that may offer organizations the opportunity to enjoy scalable, offsite, hosted environments for their enterprise data. There are three typical deployment models that exist including Public, Private and Hybrid and all of these types of environments pose significant challenges to organizations who wish to ensure the practices of defensible, timely, production of electronically stored information. The majority of digital forensic, legal, and E-Discovery experts included in this paper agree that E-

Discovery in remotely hosted social media and cloud computing environments is difficult or impossible based on current cloud provider offerings.

A number of software vendors are responding to the opportunities that these difficulties provide by offering solutions to capture social media and Web 2.0 content as it is created to solve some of the problems that exist with this unstructured remotely located information. However, the cloud providers cannot or will not allow forensic analysis in their environments due to the security risks to their other customers and the difficulties in isolating instances of organizational data on their systems.

Cloud environments spread the data over many servers, often in various parts of the world where there are vastly different national policies and laws regarding data. They may also have privileged users in their employ who can access both encrypted and unencrypted data that belong to their customers, posing significant security risks. Cloud providers may not be able to ensure that deleted data is completely removed from their widely replicated systems, and data may be distributed to their partners where they do not regulate their staff and operational policies.

The literature suggests that it would be possible to perform forensics on cloud systems. This would however require significant assistance from the cloud technical staff, and may require unacceptable security rights granted to their customers in order to complete a forensic investigation. Server logs contain co-mingled data from a number of cloud customers utilizing their systems, preventing capture and analysis of important network data. Metadata is not available in most cases and organizations have no or little control over the computing infrastructure in a cloud environment.

Recommendations based on the Literature Review Only

The literature review supports the problem that cloud computing poses unreasonable difficulties in performing E-Discovery. The practice of digital forensics and therefore defensible E-Discovery may be impossible in some cloud computing environments making it an unsuitable choice for organizations that proactively plan their governance strategies. Further statistical analysis would support this recommendation and could provide organizations with information on the cloud providers who are most responsive to E-Discovery requirements that will support their demands for transparency.

The literature included in this paper suggests that organizations considering a cloud computing environment must be diligent in investigating their options and discussing the eight cloud computing security risks. These include privileged user access, regulatory compliance, data location, data segregation, recovery, investigative support, long-term viability, and termination considerations (Brodkin, 2008).

References

Akers, S., Mason, J. K., & Mansmann, P. L. (2011, December 6). *An Intelligent Approach to E-Discovery*. Retrieved January 16, 2011, from http://www.umiacs.umd.edu/~oard/desi4/proceedings.pdf

Birk, D., & Wegener, C. (2011). *Technical Issues of Forensic Investigations in Cloud Computing Environments*. Retrieved November 15, 2011, from http://code-foundation.de/stuff/2011-birk-cloud- forensics.pdf

Brodkin, J. (2008, July 2). Gartner: Seven cloud-computing security risks: Data integrity, recovery, privacy and regulatory compliance are key issues to consider. In *Network World*. Retrieved November 21, 2011, from http://www.networkworld.com/news/2008/070208-cloud.html

Clark, J. (2011, March 18). AWS service interruptions raise doubts over reliability. In *ZDNET News and Analysis*. Retrieved November 27, 2011, from http://www.zdnet.co.uk/news/cloud/2011/03/18/aws-service-interruptions-raise-doubts-over-reliability-40092192/

Computer Crime Research Center Staff. (2011, June 7). The hidden cost of cybercrime. In *Computer Crime Research Center*. Retrieved December 7, 2011, from http://www.crime-research.org/news/07.06.2011/3873/

Cornell University Law School Staff. (n.d.). Rule 26. Duty to Disclose; General Provisions Governing Discovery. In *Cornell University Law School - Legal Information Institute*. Retrieved November 20, 2011, from http://www.law.cornell.edu/rules/frcp/Rule26.htm

Crosby, S. (2011, March 29). Why the Cloud is Actually the Safest Place for Your Data. In *Mashable Business*. Retrieved November 17, 2011, from http://mashable.com/2011/03/29/cloud-computing-security/

Delport, W., Oliver, M. S., & Kohn, M. (2011). *Isolating a Cloud Instance for a Digital Forensic Investigation*. Retrieved December 12, 2011, from http://icsa.cs.up.ac.za/issa/2011/Proceedings/Research/Delport_Olivier_Kohn.pdf

Frazier, J. (2011, November 4). The CIO's Guide to e-Discovery in the Cloud. In *CIO Update*. Retrieved December 15, 2011, from http://www.cioupdate.com/financial-strategies/the-cios-guide-to-e-discovery-in-the-cloud.html

Garfinkel, S. L. (2010). Digital forensic research: The next 10 years. *Digital Investigation, 7*, S64-S73. Retrieved December 28, 2011, from http://dfrws.org/2010/proceedings/2010-308.pdf

Golden, B. (2011, May 27). Cloud CIO: The Two Biggest Lies About Cloud Security. In *Network World*. Retrieved November 26, 2011, from http://www.networkworld.com/news/2011/052711-cloud-cio-the-two-biggest.html?page=1

Heiser, J., & Nicolett, M. (2008, June 3). Assessing the Security Risks of Cloud Computing. In *Gartner*. Retrieved November 27, 2011, from http://my.gartner.com/portal/server.pt?open=512&objID=260&mode=2&PageID=3460702&resId=685308&ref=QuickSearch&sthkw=G00157782

Hickey, A. R. (2010, April 9). Cloud Computing Security Risks Outweigh Benefits: Survey. In *CRN*. Retrieved November 26, 2011 from http://www.crn.com/news/security/224202475/cloud-computing-security-risks-outweigh-benefits-survey.htm;jsessionid=Y1-+cPO7bnMtKNf+VBhg6g**.ecappj02

Hickey, A. R. (2011, May 25). Feds Shy Away from Public Cloud, Call for Security. In *CRN*. Retrieved November 26, 2011, from http://www.crn.com/news/cloud/229625618/feds-shy-away-from-public-cloud-call-for-security.htm

Hon, K., Hornle, J., & Millard, C. (2011, September 28). Data Protection Jurisdiction and Cloud Computing. In *Social Science Research Network*. Retrieved November 27, 2011, from http://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID1950696_code1577160.pdf?abstractid=1924240&mirid=1

Hunter, S. (2011, July 14). Ascending to the Cloud Creates Negligible E-Discovery Risks. In *E-Discovery Bytes*. Retrieved December 15, 2011, from http://ediscovery.quarles.com/2011/07/articles/information-technology/ascending-to-the-cloud-creates-negligible-ediscovery-risk/?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+E-discoveryB

Inmon, W. H. & Krishnan, Krish. (2011). Building the unstructured data warehouse. [Books24x7 version] Available from http://common.books24x7.com.ezproxy1.apus.edu/toc.aspx?bookid=38016.

Jaeger, J. (2011, August). Fitting Social Media into your e-Discovery Regime. In *Compliance Week*. Retrieved January 17, 2012, from http://www.complianceweek.com/pages/login.aspx?returl=/fitting-social-media-into-your-e-discovery-regime/article/209880/&pagetypeid=28&articleid=209880&accesslevel=2&expireddays=0&accessAndPrice=0

Jaeger, J. (2011, December). Social Media Tops e-Discovery Challenges for 2011. In *allBusiness*. Retrieved January 17, 2012, from http://www.allbusiness.com/legal/trial-procedure-pretrial-discovery-electronic/15520934-1.html

James, R. (2011, December 29). Data Mapping and ESI. In *DiscoveryResources.org*. Retrieved October 9, 2011, from http://www.discoveryresources.org/technology-counsel/sound-evidence/data-mapping-esi/

Mell, P., & Grance, T. (2011, September). NIST Cloud Computing Program. In *National Institute of Standards and Technology*. Retrieved November 20, 2011, from http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf

Meyer, Esq., E. B. (2009). Largent v. Pena. In *The Employer Handbook*. Retrieved January 30, 2012, from http://www.theemployerhandbook.com/Largent.pdf

Nelson, S., Simek, J., & Foltin, J. (2009). The Legal Implications of Social Networking. In *Sensei Enterprises*. Retrieved December 28, 2011, from http://www.senseient.com/articles/pdf/The_Legal_Implications_of_Social_Networking.pdf

Oreskovic, A., & Wohl, J. (2011, April 21). Amazon cloud disruption hits some startups. In *Reuters*. Retrieved November 27, 2011, from http://www.reuters.com/article/2011/04/21/us-amazon-cloud-idUSTRE73K5X120110421

Pass, R. (2011, April 16). *e-Discovery: A View from the Bench and the Practitioner*. Retrieved January 12, 2012, from http://www.legalist.com/scottsdale2011/papers/Plenary2-2.pdf

Pelz-Sharpe, A. (2011, May 30). Pros and Cons of Content Management in the Cloud. *EContent*, 27-28. Retrieved November 27, 2011, from http://www.econtentmag.com/Articles/Column/Technology-Watch/Pros-and-Cons-of-Content-Management-in-the-Cloud-75625.htm

Reilly, D., Wren, C., & Berry, T. (2011, March). Cloud Computing: Pros and Cons for Computer Forensic Investigations. *International Journal Multimedia and Image Processing*, *1*(1). Retrieved November 26, 2011, from http://www.infonomics-society.org/IJMIP/Cloud%20Computing_Pros%20and%20Cons%20for%20Computer%20Forensic%20Investigations.pdf

Schuler, K., Peterson, C.P., Vincze, E. (2009). *E-discovery: Creating and Managing an Enterprisewide Program* (Kindle ed.). Burlington, MA: Syngress Publishing, Inc.

Stuart, E. (2011, October 28). The Value of Skype Messenger Data in eDiscovery Revealed. In *7Safe eDiscovery Services*. Retrieved January 6, 2012, from http://ediscovery.7safe.com/wp-content/uploads/2011/10/STUART-Skype-in-eDiscovery-SJC.pdf

Swindler, A. (2011, November 17). Showdown: Public vs. Private Cloud. In *Industry Week*. Retrieved November 26, 2011, from http://www.industryweek.com/articles/showdown_public_vs-_private_cloud_26042.aspx?SectionID=4

Taylor, M, J Haggerty, D Gresty, and D Lamb. (2011). "Forensic investigation of cloud computing systems." *Network Security* Mar. 2011: 4-10. Print.

Vogel, P. (2011, November 9). Cloud Computing - New Buzzword, Old Legal Issues. In *E-Commerce Times - Tech Law*. Retrieved November 26, 2011, from http://www.ecommercetimes.com/story/Cloud-Computing---New-Buzzword-Old-Legal-Issues-73714.html

Weigel, J. (2011, September 26). Cybercrime: A billion-dollar industry. In *Computer Crime Research Center*. Retrieved December 12, 2011, from http://www.crime-research.org/news/26.09.2011/3883/