



# History, Historiography and the Hermeneutics of the Hard Drive

Mark Pollitt

## ► To cite this version:

Mark Pollitt. History, Historiography and the Hermeneutics of the Hard Drive. 9th International Conference on Digital Forensics (DF), Jan 2013, Orlando, FL, United States. pp.3-17, 10.1007/978-3-642-41148-9\_1 . hal-01460612

**HAL Id: hal-01460612**

**<https://inria.hal.science/hal-01460612>**

Submitted on 7 Feb 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

## Chapter 1

# HISTORY, HISTORIOGRAPHY AND THE HERMENEUTICS OF THE HARD DRIVE

Mark Pollitt

**Abstract** This paper contrasts the traditional metaphors for digital forensics – computer science, geology and archeology – with the new metaphors of history and historiography. Narratology, the study of how narratives operate, is used to develop a construct for identifying narratives from within digital evidence. Knowledge management is suggested as a core digital forensic process. The paper describes how the investigative paradigm and traditional theories of forensic science can be integrated using two theoretical constructs, the hermeneutic and narrative theories of digital forensics. Also, natural language processing techniques are used to demonstrate how subjects can be identified from the Enron email corpus.

**Keywords:** Forensic analysis, narratology, hermeneutics, knowledge management

## 1. Introduction

Digital forensics is a science that is still very much in its formative stages. Digital forensic practitioners struggle to conduct forensic examinations of ever larger data sets that are stored and communicated with increasing technical complexity by users who integrate hardware and software, such as smart phones and web applications, more and more tightly into their daily lives. Subjectively, practitioners may sense a disconnect, perhaps even frustration, between their methodologies and their goal of identifying probative evidence as digital forensic examination and analysis become more challenging.

In the following discussion, the term “digital forensic examination” refers to the documentation, identification and extraction of data from digital evidence. The term “forensic analysis” is used to describe the review and use of the data extracted from digital evidence for legal

purposes. In other words, analysis contextualizes the extracted data into the operational environment.

This paper uses metaphors from the published literature to examine how the discipline of digital forensics has been traditionally viewed. Also, it investigates how these metaphors may limit the analytical aspects of the discipline. Additional metaphors are suggested to help illuminate the disconnect between examination and analysis. These metaphors suggest a paradigm that allows for the development of a theoretical construct for digital forensic examination and analysis. The theories that make up this new paradigm are suggested as a Kuhnian candidate paradigm for the science of digital forensics.

## 2. Traditional Forensic Science Paradigm

Inman and Rudin [10] offer one of the best known and organized paradigms for forensic science. They specify a model where there are four – and only four – forensic processes: identification, classification/individualization, association and reconstruction. The inference is that any forensic examination can be categorized into one or more of these four processes. Because, as we will shortly see, individualization is an extension of classification, we will refer to these as Inman and Rudin’s four forensic questions.

Identification is simply the scientific ability to define the nature of an object. Legally, it may be sufficient to chemically identify a controlled substance such as cocaine. The possession of cocaine is prohibited; therefore, once benzoyl-methyl-ecgonine (the chemical name for cocaine) is identified in a sample, no further qualitative examination may be necessary [10].

Classification is the ability to define objects as coming from a common origin. A bullet recovered from a homicide victim, because of its dimensions, weight, shape and markings may be identified as coming from a Smith and Wesson 9 mm pistol. Because there are millions of such pistols, this is considered “class evidence.” The gun that fired the bullet belongs to the “class” of all Smith and Wesson 9 mm pistols. If the gun that was used to fire this bullet is available, then the microscopic markings on the bore of the barrel could be used to determine that the recovered gun is, in fact, the source of the bullet recovered from the victim, to the exclusion of all other pistols. The result is that the evidence is “individualized.”

Because of this relationship between the processes of classification and individualization, we refer to them as a single forensic question. All individualizations are not binary valued – there may be levels of

certainty. Inman and Rudin [10] give a very instructive approach for dealing with the logical problem of levels of certainty given the inability to prove a hypothesis.

The last two principles described by Inman and Rudin are association and reconstruction. Association is the ability to “infer contact” between two pieces of evidence while reconstruction is the ability to order “events in the relative space and time based on the physical evidence” [10]. A fingerprint may identify the perpetrator, but when the fingerprint is found at the scene of a crime, it associates the perpetrator with the crime scene. The entry and exit wounds on a victim’s body compared with the locations of the spent bullets may allow forensic scientists to reconstruct the order of the shots and the position of the body at each impact. These two principles differ from the first two in a very significant way – they place the physical evidence in an investigative context [10].

While Inman and Rudin’s forensic questions are very useful in answering questions about physical evidence (including digital evidence), they are, in many ways, too granular to contextualize much of the data in digital evidence. The investigator and prosecutor tend to use a different set of questions. They tend to use the “who,” “what,” “when,” “where,” “why” and “how” questions. While these questions are useful for investigators and prosecutors, it is difficult to directly address them using available digital forensic tools and techniques. As a result, it is necessary to find a way to bridge the gap between the two paradigms.

Most digital forensic textbooks focus on tools and artifacts. Furthermore, the textbooks provide only generalized approaches for the actual examination and analysis of digital evidence [2–4, 11]. While digital forensic practitioners need this information, it does not help them to decide what information to collect, analyze and report.

The practitioner of traditional forensic science offers, within a discipline such as firearms examination, a limited set of examinations that answer a relatively small set of questions. For example, given a fired bullet or casing, a firearms examiner can only provide information concerning the physical characteristics of the bullet and/or casing that may answer what cartridge (e.g., 9 mm or .38 Special) or firearm (e.g., Smith and Wesson revolver or Colt pistol) was used [9]. The digital forensic examiner’s task is much more complex. The examiner must spend a great deal of time custom designing a process and the selection criteria that will recover, identify and extract only the most pertinent information for each individual case.

### 3. Engaging Other Paradigms

Farmer and Venema [8] articulate the use of a different kind of paradigm. They analogize the deleted data on a hard drive as “fossilized.” They then extrapolate digital forensics to geology and archeology. They analogize the creation of data, by the operation of the computer, to the physical forces of nature such as plate tectonics and volcanoes. They call this “digital geology.” What defines this aspect is that the artifacts are caused by the inherent operation of the computer and are not products of user action. In contrast, Farmer and Venema [8] use the term “digital archeology” to describe the artifacts of human intervention. These metaphorical approaches are helpful in differentiating examinations in which the activity is focused on a computer from those examinations involving user data. The use of the term archeology also implies the authorial nature of the data. The users are “writing” their history. This is not particularly helpful in identifying and selecting probative files or constructing knowledge from the data. It does, however, suggest an analogous approach, which is described below.

Farmer and Venema [8] describe the “discovery” of digital evidence as an archeology of system artifacts. This could be characterized as a “computer science” approach to digital forensics. By careful extraction of data and analysis, the examiner reconstructs computer activity. This approach has great merit in situations where the role of a computer in an investigation is as a victim, a weapon and sometimes as an instrumentality. Unfortunately, these represent the minority of digital forensic cases. Far more common are fraud, child pornography, intellectual property theft and forgery that comprise most personal and economic crimes. In such cases, a forensic examination must provide what can be gleaned from the contents of files and their fragments. Investigators and lawyers want the emails, memos, photographs, spreadsheets and social/personal connections that speak to the actions and intentions that comprise the case.

This suggests that Farmer and Venema’s metaphor can be extended beyond geology and archeology to the ethnography, history, literature and sociology of the computer. Forensic practitioners are interested in searching for the activities of users and the textual (in the broadest sense) record of their activities. The user activities and content of the files are mediated by cultural, social and technical factors unique to the users. Instead of merely analyzing the content of the computer hard drive to obtain the record of a machine, it is vital to conduct the process as part archaeological dig, part anthropological study of a large, semi-organized data repository, and part unedited anthology. The mission

of the digital forensic examiner is to find the data (files and fragments) that answer the relevant questions appropriate to the particular case, characters and crime. If the fossilized records of computer activity correspond to geology and archeology, then the fossilized remains of the content might be styled as history.

#### 4. History and Historiography

According to Collingwood [5], “[t]he value of history, then, is that it teaches us what man has done and thus what man is.” In many ways, what digital forensic practitioners share with investigators is the identification and documentation of what a person has done, and by extension, what he or she is. This suggests that digital forensic practitioners are to some extent historians. Collingwood recognized that history can be viewed as both science and philosophy. He also tells us that different sciences tell us different things. In the case of history, it is the “actions of human beings that have been done in the past” and history is essentially about interpreting evidence.

In order to understand what digital forensics might gain from a historical paradigm, it is instructive to consider the activities and comments of some of the most renowned historians of our time.

David McCullough is a best-selling author, whose book, *John Adams*, about America’s second President, won a Pulitzer Prize winner and was the subject of an HBO mini-series. In his 2003 Jefferson Lecture at the National Endowment of the Humanities, McCullough [13] stated: “No harm’s done to history by making it something someone would want to read.”

Another important historian is Ken Burns, who has truly made history accessible. Burns takes letters, photographs, film clips and interviews and weaves them into a “story.” He takes the data of the historical record and turns it into a narrative.

The historian Lawrence Stone [15] defines a narrative as follows: “it is organized chronologically; it is focused on a single coherent story; it is descriptive rather than analytical; it is concerned with people not abstract circumstances; and it deals with the particular and specific rather than the collective and statistical.” Indeed, Stone – and McCullough and Burns – tell us that narrative makes history accessible.

Historiography is the term for the way in which history is communicated. The selection of specific documents, photographs and historical figures mediates how history is understood. The way in which the selected historical references are presented and the way the story is told have a substantial impact on the understanding of the events that are

described. It is the historiography that makes a McCullough book or a Burns documentary powerful and compelling. The selection of the salient facts is fundamental, but their presentation ultimately defines their value.

In the same way, the selection of particular content or artifacts during a digital forensic examination is vitally important. However, if they are not presented in the context of the investigative or legal narrative, they may be overlooked or even ignored. How the content and artifacts are presented strongly influences the use of forensic data in investigations. The selection and presentation of data must create a narrative that reflects what is known about the crime, the perpetrator and the evidence. In other words, the manner in which the digital forensic examiner builds the narrative defines the hermeneutics of the hard drive.

## 5. Narrative as a Paradigm

Digital forensics is about several nested narratives. There is the narrative of the evidence itself: how was it created, in what order things were written and when. The evidence can be divided into the content authored by the user and the metadata that describes the provenance of the content. Farmer and Venema [8] might describe this as geological and archaeological narratives. Then there are the narratives told by the content. Because digital storage media is capable of multipurpose use, content is often related to a wide variety of user activities. In emails alone, it is common to find numerous discussions about disparate personal and professional genres involving a variety of individuals and groups. A hard drive is not like a novel; rather, it is part diary, part anthology and part notebook. While this makes the content complex, the goal of a digital forensic examination is to identify the pertinent data and present it in a useful manner. The problem can be framed as the need to select pertinent narratives from the evidence and assemble them into a coherent meta-narrative.

## 6. Hermeneutic Theory

Digital forensics engages two fundamental approaches: a computer science approach and an investigative approach. But a tension exists between the two approaches.

The traditional computer science forensic – or technical – approach focuses on the technical aspects of the evidence and seeks to produce reports and testimony that are scientifically defensible. In this approach, the digital forensic examiner deconstructs the physical evidence, consisting of captured and stored digital data, into a series of artifacts. The

artifacts include literal data as well as operating system, file system and network metadata. The forensic examiner seeks, through the forensic examination process, to answer the four questions described by Inman and Rudin [10] in their forensic taxonomy. Forensic examiners produce forensic reports and testimony that can be described as a meta-narrative combining the forensic process undertaken in the current examination, the operation of the technologies associated with the artifacts identified, and the artifacts themselves. These artifacts may be selected because they are probative in either the case or the legal narrative, but the examiner's forensic interest is in their presence and provenance, not in the content *per se*.

In contrast, the investigative approach seeks to discover the people and the events that constitute proof of a crime or tort. This is often accomplished using a different subset of the evidence to answer different questions. While the technical information and metadata obtained from a forensic examination may be used in investigative analysis, it is principally the content of the files and fragments that form the majority of the analysis. This does not imply that the products of the technical examination are not useful or not commonly used in the analysis. These, in large measure, form the skeleton of the analysis, but it is the content of the artifacts that provides most of the material from which the investigative narratives are created. The person with the investigative/analytical role, regardless of whether he is an attorney, investigator or forensic examiner, seeks to construct a meta-narrative that comprises two main elements, the case narrative and the legal narrative. The former is the narrative constructed by answering the investigative questions of who, what, when, where, why and how? This case narrative takes the answers to the investigative questions and instantiates them into the elements of the crime or tort, in the process creating the legal narrative. The evidence utilized to document these elements are the physical evidence (which includes the digital evidence) and the testimony (both lay and expert testimony).

Despite any notions to the contrary, these two approaches have been interrelated since the very first digital forensic examination. The relationships have changed as technology has advanced, laws have adapted and processes have evolved. In the earliest days of digital forensics, it was the investigative approach that drove the process. As practitioners gained experience, as the courts began to admit digital evidence and as the technology grew more complex, the technical approach became more important. In the mid-1990s, technology was advancing rapidly, but the ordinary citizen had little technical knowledge. To prove the reliability



*Table 1.* Comparison of technical and investigative approaches.

	<b>Technical Approach</b>	<b>Investigative Analysis Approach</b>
Evidence	Artifacts (operating system, file system, application metadata, file contents)	Content and communications from recovered artifacts
Process	Forensic examination	Investigative analysis
Context	Information technology systems	Case and legal context
Answers	Forensic questions	Investigative questions
Explicative Approach	Meta-narrative of forensic process, technology and evidentiary artifacts	Meta-narrative of case narrative and legal elements

of digital evidence, practitioners relied more and more on the technical aspects to prove the legitimacy of their proffered evidence.

Technical examinations were not, however, answering the investigative questions. Many investigators sought to perform the analysis part of the process and rely on technicians merely to provide reliable data. As digital evidence data sets grew ever larger, the ability of investigators to conduct efficient analyses diminished. The increased volume of evidence transformed the problem to one of knowledge management. The use of technologies, such as natural language processing and XML, may allow for technically-assisted knowledge management. But these technologies will not, by themselves, provide much assistance. The transformation of data into information and the transformation of information into knowledge are cognitive processes. While tools can help present data or information to analysts, they do not, on their own, produce information or knowledge. In order to transform data to information, it is necessary to examine the content and to situate it in an investigative or forensic context. As a result, it is vital to move beyond tools that focus on the technology and apply content-focused methodologies that utilize technologies to assist analysts in contextualizing data and information.

Both the technical and investigative approaches are knowledge management processes and, as such, must add value to the data. Collectively, the two approaches can be viewed as overarching theoretical constructs for digital forensics. Digital forensics cannot exist without both these approaches and, thus, the two approaches comprise the core paradigm of the science of digital forensics. Table 1 summarizes the key elements of the two approaches.

## 7. Hermeneutic Theory of Digital Evidence

Based on the preceding discussion, we suggest the following formal paradigm – or theory – for digital forensics:

- The legal system utilizes data stored or transmitted in digital form as evidence.
- The digital evidence must meet the reliability and authenticity tests required by the legal system.
- Forensic examinations preserve the integrity of the original evidence by technical means.
- The products of the examination process consist of artifacts, which can be sub-divided into metadata and content.
- The products of the digital forensic process are utilized to answer two interrelated sets of questions: forensic questions and the investigative questions.
- The digital forensic examination process focuses on answering the forensic questions.
- The digital forensic analysis process utilizes the products from the digital forensic examination process and generally focuses on answering the investigative questions.
- The examination and analysis phases of the digital forensic process are knowledge management processes. These processes should add value to the data and information developed during each phase.
- The forensic and investigative processing of evidence results in one or more meta-narratives, which are based on a combination of technical artifacts, metadata and content. These results form a synthesis of the forensic and investigative processes.
- The goal of the digital forensic process is to provide knowledge in the form of actionable intelligence, investigative leads, testimony and/or probative evidence.

## 8. Narrative Theory

The last portion of the hermeneutic theoretical construct states, in effect, that all digital evidence is – at one or more levels – a narrative. These narratives contribute to and are parts of the technical and the investigative meta-narratives. The construct also states that, in order

to create meta-narratives, content is a key element of the examination and analysis processes. Implicit in this notion is that content contains a narrative, contributes to a narrative or is in and of itself a narrative.

Since a narrative is so important to the analysis of evidence, how should we seek this crucial element? One solution is to utilize the core concepts of narratology. The field of narratology seeks to understand what constitutes a narrative, how it is structured and how it works. Bal [1], one of the pioneers of narratology, argues that there are effectively three elements that define a narrative: chronology/logic, event/action and actors.

Other researchers [6, 7, 16] suggest that it may be possible to search for narratives by utilizing an approach that attempts to identify narratives by utilizing the semantics of the content. Because this is primarily a semantic analysis, the use of sentences as a basic unit of analysis is an appropriate first approach.

The goal of a semantic analysis is to develop knowledge of subjects/actors, actions/events and chronology. Therefore, a suitable approach is to apply natural language processing to identify and recognize relationships among these elements. Since nouns and verbs are rough proxies for subjects/actors and actions/events, respectively, techniques that focus on these grammatical structures can improve the identification, efficiency and comprehension of narratives.

Beyond the identification of narratives, it is essential, given the vast quantities of digital evidence, to be able to identify particular narratives that are probative. At present, there do not seem to be technologies that can, by themselves, accurately evaluate the probative value of any given narrative. However, the ability of human analysts to evaluate data is limited by things like attention, fatigue and preconceptions about the data. It would, therefore, be useful to develop approaches that simultaneously reduce the volume of data that needs to be reviewed by analysts and enhance the ability of analysts to identify probative data. Reading a half million emails is not practical. Reducing this number and simultaneously increasing the likelihood that what is read is probative should be a goal of knowledge management in the digital forensic context.

These aspects of the analysis of digital evidence suggest the following hypothesis:

- **Hypothesis 1:** The identification of narratives by automated means will contribute to the efficiency and effectiveness of a forensic examiner or investigative analyst.

A second dependent hypothesis is:

- **Hypothesis 2:** Any automated process that improves the ability of a forensic examiner or investigative analyst to quickly identify probative narratives will improve the efficiency and effectiveness of the process.

Two other hypotheses are:

- **Hypothesis 3:** The use of nouns and verbs will assist in the identification of general and probative narratives more economically than reading complete texts.
- **Hypothesis 4:** Natural language processing software can assist in the identification of probative narratives by the use of lexical, grammatical and semantic techniques.

## 9. Identifying Narrative Elements

Todorov and Weinstein [16] identify the elements of a narrative as: (i) the subject, identified as a noun; (ii) the predicate, which is “always” a verb; and (iii) the adjective, which infuses a “quality” without changing the situation. After describing a grammatical analysis of structure, Todorov and Weinstein suggest that this grammatical approach can be used to further the study of narrative syntax, theme and rhetoric. If a corpus of textual digital evidence were to be processed using a “named entity” extraction technique, then it might be possible to identify the subjects of the narratives contained in the corpus.

A series of experiments were conducted to test this approach. In order to provide an investigative focus for the experiments, it was decided to study a specific and well-documented fraudulent scheme perpetrated by a group of Enron employees. The scheme involved the fraudulent raising of the price of electricity purchased by California power companies. This scheme was described by McLean and Elkind [14] in a book entitled *The Smartest Guys in the Room: The Amazing Rise and Scandalous Fall of Enron*.

In order to test the utility of natural language processing in identifying the subjects of this scheme, a subset of the Enron corpus comprising 2,385 emails from user-labeled directories involving the word “California” were processed for named entities. These emails resulted in more than 11,000 uniquely-named entities. The same natural language processing Python script was used to process Chapter 17 (*Gaming California*) in McLean and Elkind’s book [14], which discussed the fraudulent scheme.

Table 2. Results of named entity extraction.

Source	Uniquely-Named Entities
McLean and Elkind Chapter 17	151
Enron Employee California Emails (n = 2,385)	11,336
McLean and Elkind Chapter 17 and Emails	98

Table 2 shows the results of the experiment. The 98 uniquely-named entities, which were in the book chapter and in the emails, represent 64.9% of the 151 uniquely-named entities in the book chapter. In other words, a little less than 65% of the named entities found in the book chapter were also found in the emails involving the term “California.” This ratio, which we call the “commonality ratio,” is given by:

$$C = \frac{NI^k}{NE^k \cap NE^q}$$

where  $NE^k$  is the set of named entities in the known text and  $NE^q$  is the set of named entities located in the questioned text.

Table 3. Results of named entity extraction from Chapter 17 and emails.

	Chap. 17 Only	Chap. 17 and Emails	Commonality Ratio
Original Results	151	98	64.9%
Misclassified and Misspelled Removed	143	98	68.5%
External Literary References Removed	139	98	70.5%

A review of the results, while promising, raised some concerns about the “cleanliness” of the data. A number of the named entities had been misclassified by the software and were not actually named entities. Other words were misspelled and were incorrectly classified. Also, a number of correctly-identified named entities, primarily in the book chapter, were names of journalistic sources. In an effort to clean the data, all the misclassified and misspelled words and journalistic references were removed. Table 3 shows the results.

In an effort to determine if the approach is capable of discriminating between different sets of narratives, a second experiment was undertaken. Another chapter from McLean and Elkind’s book, which does not focus on the California scheme, but on an accounting fraud masterminded by Andrew Fastow, was used. This chapter, entitled *Andy Fastow’s Secrets* (Chapter 11), was processed in the same manner as Chapter 17. The

Table 4. Results of named entity extraction from Chapter 11 and emails.

	Chap. 11 Only	Chap. 11 and Emails	Commonality Ratio
Original Results	157	62	39.5%
Misclassified and Misspelled Removed	153	62	40.5%
External Literary References Removed	152	62	40.8%

data was also cleaned in the same fashion as in the previous experiment. The results are shown in Table 4.

Given the large number of emails that contained the names of people and organizations that were pertinent to a wide range of Enron operations, it is remarkable that the commonality ratios are so different. This is further reinforced by the fact that several major actors in the Enron saga are named in both chapters, thus increasing the ratio for both sets. Even in the worst case scenarios – where the minimum commonality between the California emails and Chapter 17 is compared with the maximal commonality between the California emails and Chapter 11 – the difference is 62.8%. This would appear to be a significant level of differentiation between two narratives that share a large number of common subjects.

## 10. Conclusions

Digital forensics is a nascent science that lacks a substantial theoretical foundation. Most of its existent theory is borrowed from computer science or forensic science. This paper has proposed two theoretical constructs, the hermeneutic and narrative theories of digital forensics, that attempt to integrate the scientific and investigative aspects of digital forensics. Narrative theory seeks to reify the fundamentally textual nature of digital evidence in the specific genre of the narrative. Proffering such theoretical constructs may be viewed as ambitious, perhaps even presumptuous, but they are offered in the spirit of research. The constructs can be evaluated, criticized, tested, disproved and improved. Any discipline that is deemed “scientific” as specified by Kuhn must have an underlying paradigm. The theoretical constructs are offered as a “shared example” or “candidate paradigm” in the sense of Kuhn [12]:

History suggests that the road to a firm research consensus is extraordinarily arduous. In the absence of some candidate for paradigm, all the facts that could possibly pertain to the development of a given science are likely to seem equally relevant. Only very occasionally, as in the case of ancient statics, dynamics, and geometrical optics, do facts collected

with so little guidance from pre-established theory speak with sufficient clarity to permit the emergence of a first paradigm.

The use of natural language processing to demonstrate the potential for the use of narrative shows promise. While the experiments conducted were neither elegant nor definitive, they suggest that combining narratology and natural language processing have promise and are worthy of further research.

## References

- [1] M. Bal, *Narratology: Introduction to the Theory of Narrative*, University of Toronto Press, Toronto, Canada, 2009.
- [2] B. Carrier, *File System Forensic Analysis*, Pearson Education, Upper Saddle River, New Jersey, 2005.
- [3] H. Carvey, *Windows Forensic Analysis DVD Toolkit*, Syngress, Burlington, Massachusetts, 2007.
- [4] E. Casey, *Digital Evidence and Computer Crime: Forensic Science, Computers and the Internet*, Academic Press, Waltham, Massachusetts, 2011.
- [5] R. Collingwood, *The Idea of History*, Oxford University Press, Oxford, United Kingdom, 1994.
- [6] A. de Waal, J. Venter and E. Barnard, Applying topic modeling to forensic data, in *Advances in Digital Forensics IV*, I. Ray and S. Shenoi (Eds.), Springer, Boston, Massachusetts, pp. 115–126, 2008.
- [7] J. Doyle, Mapping the World of Consumption: Computational Linguistics Analysis of the Google Text Corpus, Working Paper, Cardiff Business School, Cardiff University, Cardiff, United Kingdom, 2010.
- [8] D. Farmer and W. Venema, *Forensic Discovery*, Addison-Wesley, Upper Saddle River, New Jersey, 2005.
- [9] E. Hueske, Firearms and tool marks, in *The Forensic Laboratory Handbook Procedures and Practice*, A. Mozayani and C. Noziglia (Eds.), Humana Press, Totowa, New Jersey, pp. 143–176, 2006.
- [10] K. Inman and N. Rudin, *Principles and Practice of Criminalistics: The Profession of Forensic Science*, CRC Press, Boca Raton, Florida, 2000.
- [11] W. Kruse and J. Heiser, *Computer Forensics: Incident Response Essentials*, Pearson Education, Indianapolis, Indiana, 2002.
- [12] T. Kuhn, *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago, Illinois, 1996.

- [13] D. McCullough, The Course of Human Events, David McCullough Lecture, National Endowment for the Humanities, Washington, DC ([www.neh.gov/about/awards/jefferson-lecture/david-mccullough-lecture](http://www.neh.gov/about/awards/jefferson-lecture/david-mccullough-lecture)), 2003.
- [14] B. McLean and P. Elkind, *The Smartest Guys in the Room: The Amazing Rise and Scandalous Fall of Enron*, Portfolio, New York, 2003.
- [15] L. Stone, The revival of narrative: Reflections on a new old history, *Past and Present*, no. 85, pp. 3–24, 1979.
- [16] T. Todorov and A. Weinstein, Structural analysis of narrative, *Novel: A Forum on Fiction*, vol. 3(1), pp. 70–76, 1969.