



WILEY CORPORATE F&A

CYBER FORENSICS

From Data
to Digital Evidence



ALBERT J. MARCELLA, JR.
FREDERIC GUILLOSSOU

Cyber Forensics

From Data to Digital Evidence

ALBERT J. MARCELLA, JR., PhD, CISA, CISM
FREDERIC GUILLOSSOU, CISSP, CCE



WILEY

John Wiley & Sons, Inc.

Copyright © 2012 by John Wiley & Sons, Inc. All rights reserved.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 646-8600, or on the Web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permissions>.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books. For more information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

Marcella, Albert J.

Cyber forensics : from data to digital evidence / Albert J. Marcella, PhD, CISA, CISM, Frederic Guillossou, CISSP, CCE.

pages cm.— (The Wiley Corporate F&A series)

Includes index.

ISBN 978-1-118-27366-1 (hardback); ISBN 978-1-118-28268-7 (ebk);

ISBN 978-1-118-28505-3 (ebk); ISBN 978-1-118-28731-6 (ebk)

1. Forensic sciences—Technological innovations. 2. Electronic evidence.

3. Evidence, Criminal. 4. Criminal investigation. 5. Computer crimes—Investigation. I. Guillossou, Frederic, 1970 - II. Title.

HV8073.5.M168 2012

363.250285—dc23

2011048568

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Al Marcella

To my wife

Diane . . .

A sunbeam to warm you,

A moonbeam to charm you,

A sheltering angel, so nothing can harm you.

*May you always know how happy you make me, and how much
I love you! Love doesn't make the world go round; love is what
makes the ride worthwhile.*

Thank you for sharing with me, the ride of a lifetime.

$\infty + 1$

Fredric Guillosoou

To my wife and daughter

Alexandra and Nathalie

*The happy memories of the past, the joyful moments
of the present, and the hope and promise of the future.*

Preface

THE ROLE AND RESPONSIBILITY of a cyber forensic investigator is to accurately report upon actions taken to expertly identify, extract, and analyze those data that will ultimately represent evidential matter as part of an investigation of an individual who is suspected of engaging in unauthorized activities.

As an expert, a cyber forensic investigator who heavily relies upon the automated, generated results of a forensic software tool, without an intimate knowledge of how the results have been achieved, is risking not only his or her professional reputation but also the potential of a successful outcome to an investigation.

Data, the primordial building blocks of information as we know it, begins life as nothing more than electrical impulses representing an existence or lack thereof, of an electrical charge. Knowing just how these pulses end up as data, and how these data then end up as potential evidence, is an essential skill for a cyber forensic investigator.

The evolution of bits and bytes into data and finally into human-understandable text is not rocket science; somewhat technical yes, but not beyond the reach or understanding of the professional looking to gain a greater understanding of HOW data become digital forensic evidence, WHERE to look for this evidence, buried beneath hundreds of millions of bytes of data, and WHY specific data may lead the investigator to the proverbial “smoking gun.”

In communicating the results of a cyber forensic investigation, responding to the question “How did you identify the specific data you examined to reach your conclusion?” by eluding to your use of a specific cyber forensic tool without a thorough understanding of how that tool “achieved” its answer, could be professionally dangerous.

Reliance on the software to produce an answer, without a solid understanding of the HOWs, WHATs, WHYs, and the theory and logic behind *how* the answer was attained is akin to submitting all of the correct answers to a mathematics exam and failing, because you did not show your work. Knowing the answer

without knowing how you achieved the answer or how to explain how the answer was achieved is having only half of a solution.

The book you are about to read will provide you with the specific knowledge to speak confidently about the validity of the data identified, accessed, and analyzed as part of a comprehensive cyber forensic investigation.

We start small, in fact very small . . . bits and bytes small, explaining the origins of data and progressing onward, addressing concepts related to data storage, boot records, partitions, volumes, and file systems, and how each of these are interrelated and essential in a cyber forensic investigation. The role each plays in an investigation and what type of evidential data may be identified within each of these areas.

Also addressed are two often overlooked topics which impact almost every cyber-based investigation: endianness and time. Each of these topics rightly deserve their own chapter and are discussed in-depth with respect to their impact and influence on data and ultimately on the identification of digital evidence.

In an effort to more effectively introduce specific information technology (IT) and cyber forensic concepts and discuss critical cyber forensic processes, we proudly introduce Ronelle Sawyer and Jose McCarthy, employees who become involved in the theft of intellectual property.

Ronelle and Jose's activities and actions are discussed throughout the book as an ongoing case, designed to provide the reader with specific examples of the application of the cyber forensic concepts discussed throughout the 12 primary chapters of this book. Although the case and characters are fictitious, the scenario presented is not.

Along with this case, we have developed and present an exemplar forensic investigation report (Forensic Investigations, ABC Inc.), which appears as an Appendix to this book. This exemplar report provides the reader with a basic forensic report template, which summarizes the forensic investigation and case data as it would be compiled for submission to a respective authorized recipient. We realize that there are many varied ways in which the results of an investigation may be compiled and presented; the report included herein is an example of one such way.

While each investigation is unique, there will be similarities and as each case is unique on to itself, a generalized investigation approach can be constructed. We have provided you, the reader, with generalized Investigative Smart Practices (ISPs) as you hone and develop your individualistic investigative processes. These are not "best practices," but "smart practices" steps, procedures, and actions, which in general, can be applied to most cyber forensic case/investigations.

It would be illogical to try to present an investigative procedure or methodology and claim that it is universal, that it can be applied in all instances under all circumstances. As such, our ISPs cast the widest net and are applicable to most general investigative cases. It is up to you the reader to add to this base, adding specific, specialized company, department, or agency steps and procedures, which will result in a uniquely identifiable case-by-case investigative process.

Regardless of your confidence in the data identified via your investigative efforts or through the use of any specific or generalized cyber forensic software, take to heart the Russian proverb, “doveryai, no proveryai,” made famous by the late Ronald Reagan: “trust but verify!”

This book will provide you with a comprehensive examination and discussion of the science of cyber forensic investigations, what is happening behind the scenes to data and why, what to look for and where to find it . . . progressing logically, from data to digital evidence.

Al Marcella and Fred Guilloso

Acknowledgments

AS AUTHORS, LET'S be frank: It is almost impossible to be fully honest when assessing one's own work. It's also impossible to be fully independent or even neutral when attempting to assess or evaluate what one has written, no matter how hard one tries.

Thus, to remedy this truism, we, as most dedicated authors, reach out to colleagues, peers, and sometimes even to strangers (well, the publisher does) to provide us with a truly independent assessment and review of what we have written.

This assessment can occur at various stages of the development of a book, such as the one you are about to read, in segments or chapters, during its formative development stages, as a completed, draft manuscript or even once the last keystroke has been struck and development is finalized.

To achieve this sought after assessment, we have reached out to individuals whom we respect, asking them to critically review our work and to provide us with the benefit of their expertise and extensive knowledge in the fields of cyber forensics, audit, information technologies, e-discovery, and investigative sciences, as they critiqued the book you are about to read.

We are thankful for their assessment and suggestions for improvement, as they have provided us with valuable insights into refining our text and providing you the reader, with the most accurate and technically current material related to the emerging and evolving field of cyber forensic investigation and analysis.

While it is not possible to individually acknowledge all of the reviewers who have assessed our work, as some will forever remain anonymous, the authors would like to personally thank the following individuals for their insights, time, and involvement in making our development efforts result in a better overall examination and presentation of the science of cyber forensics.

To the following professionals, we say a heartfelt thank you . . .

Don Caniglia, CEGIT, CISA, CISM, FLMI
President
IT Risk Management Services, LLC

Richard J. Dippel, JD, MBA, CPA
Assistant Professor of Accounting
George Herbert Walker School of Business & Technology
Webster University

Linda C. Ertel, CISA
Security Compliance Analyst
Independent Reviewer

Steve Grimm
Webster Groves Police Department Detective
The Greater St. Louis Regional Computer Crime Education and
Enforcement Group

Detective Andy Hrenak, CFCE/A+/ACE/DFCB
Hazelwood Police Department
RCCEEG Forensic Examiner

Jeff Lukins, CISSP, CEH, MCSE, MSE
Deputy IT Sec. Mgr., NASA MITS
Dynetics Technical Services, Inc.

Doug Menendez, CISA, CIA
Audit Manager
Graybar Electric Company

Bruce Monahan, CIA, CISA, CFE, CPCU
Chief Audit Executive
Selective Insurance Group, Inc.

Greg Strauss, CCE
Computer Forensics Expert
Independent Reviewer

Although not reviewers, we also wish to thank Ronelle and Jose, for providing us with a more personal means by which we were able to convey technical, cyber forensic concepts through a realistic case example. Thank you both!

Sincerely,

Al Marcella, Ph.D., CISA, CISM

Frederic Guillosoou, CISSP, CCE

Contents

Preface	xiii
Acknowledgments	xvii
Chapter 1: The Fundamentals of Data	1
Base 2 Numbering System: Binary and Character Encoding	2
Communication in a Two-State Universe	3
Electricity and Magnetism	3
Building Blocks: The Origins of Data	4
Growing the Building Blocks of Data	5
Moving Beyond Base 2	7
American Standard Code for Information Interchange	7
Character Codes: The Basis for Processing Textual Data	10
Extended ASCII and Unicode	10
Summary	12
Notes	13
Chapter 2: Binary to Decimal	15
American Standard Code for Information Interchange	16
Computer as a Calculator	16
Why Is This Important in Forensics?	18
Data Representation	18
Converting Binary to Decimal	19
Conversion Analysis	20
A Forensic Case Example: An Application of the Math	20
Decimal to Binary: Recap for Review	22
Summary	23
Chapter 3: The Power of HEX: Finding Slivers of Data	25
What the HEX?	26
Bits and Bytes and Nibbles	27

Nibbles and Bits	29
Binary to HEX Conversion	30
Binary (HEX) Editor	34
The Needle within the Haystack	39
Summary	41
Notes	42
Chapter 4: Files	43
Opening	44
Files, File Structures, and File Formats	44
File Extensions	45
Changing a File's Extension to Evade Detection	47
Files and the HEX Editor	53
File Signature	55
ASCII Is Not Text or HEX	57
Value of File Signatures	58
Complex Files: Compound, Compressed, and Encrypted Files	59
Why Do Compound Files Exist?	60
Compressed Files	61
Forensics and Encrypted Files	64
The Structure of Ciphers	65
Summary	66
Notes	67
Appendix 4A: Common File Extensions	68
Appendix 4B: File Signature Database	73
Appendix 4C: Magic Number Definition	77
Appendix 4D: Compound Document Header	79
Chapter 5: The Boot Process and the Master Boot Record (MBR)	85
Booting Up	87
Primary Functions of the Boot Process	87
Forensic Imaging and Evidence Collection	90
Summarizing the BIOS	92
BIOS Setup Utility: Step by Step	92
The Master Boot Record (MBR)	96
Partition Table	102
Hard Disk Partition	103
Summary	110
Notes	111

Chapter 6: Endianness and the Partition Table	113
The Flavor of Endianness	114
Endianness	116
The Origins of Endian	117
Partition Table within the Master Boot Record	117
Summary	125
Notes	127
 Chapter 7: Volume versus Partition	 129
Tech Review	130
Cylinder, Head, Sector, and Logical Block Addressing	132
Volumes and Partitions	138
Summary	142
Notes	144
 Chapter 8: File Systems—FAT 12/16	 145
Tech Review	145
File Systems	147
Metadata	149
File Allocation Table (FAT) File System	153
Slack	157
HEX Review Note	160
Directory Entries	161
File Allocation Table (FAT)	163
How Is Cluster Size Determined?	167
Expanded Cluster Size	169
Directory Entries and the FAT	170
FAT Filing System Limitations	174
Directory Entry Limitations	176
Summary	177
Appendix 8A: Partition Table Fields	179
Appendix 8B: File Allocation Table Values	180
Appendix 8C: Directory Entry Byte Offset Description	181
Appendix 8D: FAT 12/16 Byte Offset Values	182
Appendix 8E: FAT 32 Byte Offset Values	184
Appendix 8F: The Power of 2	186
 Chapter 9: File Systems—NTFS and Beyond	 189
New Technology File System	189
Partition Boot Record	190

Master File Table	191
NTFS Summary	195
exFAT	196
Alternative Filing System Concepts	196
Summary	203
Notes	204
Appendix 9A: Common NTFS System Defined Attributes	205
Chapter 10: Cyber Forensics: Investigative Smart Practices	207
The Forensic Process	209
Forensic Investigative Smart Practices	211
Step 1: The Initial Contact, the Request	211
Step 2: Evidence Handling	216
Step 3: Acquisition of Evidence	221
Step 4: Data Preparation	229
Time	238
Summary	239
Note	240
Chapter 11: Time and Forensics	241
What Is Time?	241
Network Time Protocol	243
Timestamp Data	244
Keeping Track of Time	245
Clock Models and Time Bounding: The Foundations of Forensic Time	247
MS-DOS 32-Bit Timestamp: Date and Time	248
Date Determination	250
Time Determination	254
Time Inaccuracy	258
Summary	259
Notes	260
Chapter 12: Investigation: Incident Closure	263
Forensic Investigative Smart Practices	264
Step 5: Investigation (Continued)	264
Step 6: Communicate Findings	265
Characteristics of a Good Cyber Forensic Report	266
Report Contents	268
Step 7: Retention and Curation of Evidence	269
Step 8: Investigation Wrap-Up and Conclusion	273

Investigator's Role as an Expert Witness	273
Summary	279
Notes	280
Chapter 13: A Cyber Forensic Process Summary	283
Binary	284
Binary—Decimal—ASCII	285
Data Versus Code	287
HEX	288
From Raw Data to Files	288
Accessing Files	289
Endianness	290
Partitions	291
File Systems	291
Time	292
The Investigation Process	292
Summary	295
Appendix: Forensic Investigations, ABC Inc.	297
Glossary	303
About the Authors	327
Index	329

The Fundamentals of Data

THIS BOOK IS DESIGNED to address the fundamental concepts found in the emerging and rapidly evolving field of cyber forensics.

Before one can profess to be knowledgeable and fully cognizant of the breadth encompassing the professional discipline of cyber forensics, a foundation, rooted in the basics of information technology, data storage, handling, and processing, as well as how data is moved and manipulated, is essential.

For the cyber forensic investigator, data is evidence. Understanding how evidence emerges from data is pivotal; however, more important is being able to confidently articulate how evidential data was identified, collected, and processed.

As a cyber forensic investigator, simply pressing buttons or checking off options in a forensic software suite, without the knowledge of what is happening behind the scenes, creates a potential liability. Understanding the “life cycle” of data is pivotal, from its humble beginnings as electronic *bits*, evolving into bytes,

characters, then words, finally emerging as a language, as information, and perhaps eventually as evidence.

This book will provide a platform for both broadening as well as enhancing your skills in the basic elements of information technology as the technology supports and is embedded within the science of cyber forensic investigations.

As you read this book, you will encounter words that have been *italicized*. These words represent key concepts and are more fully defined by a working definition, which is included within a glossary at the end of the book. Should you desire an explanation of any *italicized* word, please refer to this glossary.

As with most tasks, one must crawl prior to walking and certainly before dashing off in a full run. Therefore, our first chapter begins naturally, at the beginning, with a discussion of the prime building blocks of data and how as a society we carbon-based humans have learned to communicate with a silicon-based technology—computers.



BASE 2 NUMBERING SYSTEM: BINARY AND CHARACTER ENCODING

Modern humans use character sets (or alphabets) to represent written sounds and words. In many alphabets, including Latin-based alphabets, each symbol or letter has its own phonetic sound.

The letter (or combination of letters, such as “ph”) is paired to its corresponding sound, forming a character code. It is through the combination of these symbols or letters that humans generate words, then phrases, and ultimately complex communication.

Symbolic characters, such as alphanumeric symbols found in Latin-based languages, work reasonably well for the complex computing power of the human brain. Computers, however, have yet to evolve to a level capable of exactly duplicating the complex processing—consistently, seamlessly, and reliably—of the human brain. Currently, computers can best communicate with other computers, in a manner based upon the principles of fundamental mathematics. Computer-to-human communication, while having evolved to a certain degree of voice replication, is still based, again, upon the principles of fundamental mathematics.

The current methodology for digital data transfer is called binary, and it is the basis for all computing technology. In order to understand how computers handle, move, store, access, save, delete, or otherwise manipulate data, it is essential to first understand the concepts of the *binary system*.