# Digital Forensics with Open Source Tools

# Digital Forensics with Open Source Tools

**Cory Altheide**

**Harlan Carvey**

Technical Editor

**Ray Davidson**

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Syngress is an imprint of Elsevier

SYNGRESS

**Acquiring Editor:** Angelina Ward
**Development Editor:** Heather Scherer
**Project Manager:** Andre Cuello
**Designer:** Joanne Blank

*Syngress* is an imprint of Elsevier
225 Wyman Street, Waltham, MA 02451, USA



**Notices**
Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods or professional practices, may become necessary. Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information or methods described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

**Library of Congress Cataloging-in-Publication Data**
Application submitted

**British Library Cataloguing-in-Publication Data**
A catalogue record for this book is available from the British Library.

ISBN: 978-1-59749-586-8

Printed in the United States of America

11 12 13 14 10 9 8 7 6 5 4 3 2 1

Typeset by: diacriTech, India



For information on all Syngress publications visit our website at *www.syngress.com*

# Contents

# About the Authors

**Cory Altheide** is a security engineer at Google, focused on forensics and incident response. Prior to Google, Cory was a principal consultant with MANDIANT, an information security consulting firm that works with the Fortune 500, the defense industrial base, and banks of the world to secure their networks and combat cyber crime. In this role he responded to numerous incidents for a variety of clients in addition to developing and delivering training to corporate and law enforcement customers.

Cory also worked as the senior network forensics specialist in the National Nuclear Security Administration's Information Assurance Response Center (NNSA IARC). In this capacity he analyzed potentially hostile code, performed wireless assessments of Department of Energy facilities, and researched new forensic techniques. He also developed and presented hands-on forensics training for various DoE entities and worked closely with members of the Southern Nevada Cyber Crimes Task Force to develop their skills in examining less common digital media.

Cory has authored several papers for the computer forensics journal *Digital Investigation* and was a contributing author for *UNIX and Linux Forensic Analysis* (2008) and *The Handbook of Digital Forensics and Investigation* (2010). Additionally, Cory is a recurring member of the program committee of the Digital Forensics Research Workshop.

**Harlan Carvey** (CISSP) is a vice president of Advanced Security Projects with Terremark Worldwide, Inc. Terremark is a leading global provider of IT infrastructure and "cloud computing" services based in Miami, Florida. Harlan is a key contributor to the Engagement Services practice, providing disk forensics analysis, consulting, and training services to both internal and external customers. Harlan has provided forensic analysis services for the hospitality industry and financial institutions, as well as federal government and law enforcement agencies. Harlan's primary areas of interest include research and development of novel analysis solutions, with a focus on Windows platforms. Harlan holds a bachelor's degree in electrical engineering from the Virginia Military Institute and a master's degree in the same discipline from the Naval Postgraduate School. Harlan resides in Northern Virginia with his family.

# Acknowledgments


## Cory Altheide

First off I want to thank Harlan Carvey. In addition to serving as my coauthor and sounding board, he has been a good friend and colleague for many years. He has proven to be one of the most consistently knowledgeable and helpful individuals I have met in the field. Harlan, thanks again for adding your considerable expertise to the book and for never failing to buy me a beer every time I see you.

I also thank Ray Davidson for his work as technical editor. His early insights and commentary helped focus the book and made me target my subsequent writing on the intended audience.

Tremendous thanks go out to the "usual suspects" that make the open source forensics world the wonderful place it is. First, thank you to Wietse Venema and Dan Farmer for creating open source forensics with "The Coroner's Toolkit." Thanks to Brian Carrier for picking up where they left off and carrying the torch to this day. Simson Garfinkel, you have my gratitude for providing the invaluable resource that is the Digital Forensics Corpora. Special thanks to Eoghan Casey, who first encouraged me to share my knowledge with the community many years ago.

To my parents, Steve and Jeanine Altheide, thank you for buying my first Commodore-64 (and the second… and the third). Thanks to my brother Jeremy Altheide and the Old Heathen Brewing Company for producing some of the finest beers around… someday.

I express infinite gratitude to my incredible wife Jamie Altheide for her never-ending patience, love, and support during the research and writing of this book. Finally, I thank my daughters Winter and Lily for reminding me every day that I will never have all the answers, and that's okay.

## Harlan Carvey

I begin by thanking God for the many blessings He's given me in my life, the first of which has been my family. I try to thank Him daily, but I find myself thinking that that's not nearly enough. A man's achievements are often not his alone, and in my heart, being able to write books like this is a gift and a blessing in many ways.

I thank my true love and the light of my life, Terri, and my stepdaughter, Kylie. Both of these wonderful ladies have put up with my antics yet again (intently staring off into space, scribbling in the air, and, of course, my excellent imitations taken from some of the movies we've seen), and I thank you both as much for your patience as for being there for me when I turned away from the keyboard. It can't be easy to have a nerd like me in your life, but I do thank you both for the opportunity to "put pen to paper" and get all of this stuff out of my head. Yes, that was a John Byrne reference.

Finally, whenever you meet Cory, give him a thundering round of applause. This book was his idea, and he graciously asked me to assist. I, of course, jumped at the chance to work with him again. Thanks, Cory.

# Introduction

## INTENDED AUDIENCE

When writing a technical book, one of the first questions the authors must answer is "Who is your audience?" The authors must then keep this question in mind at all times when writing. While it is hoped that this book is useful to everyone that reads it, the intended audience is primarily two groups.

The first group is new forensic practitioners. This could range from students who are brand new to the world of digital forensics, to active practitioners that are still early in their careers, to seasoned system administrators looking to make a career change. While this book is not a singular, complete compendium of all the forensic knowledge you will need to be successful, it is, hopefully, enough to get you started.

The second audience is experienced digital forensics practitioners new to open source tools. This is a fairly large audience, as commercial, proprietary tools have had a nearly exhaustive hold on working forensic examiners. Many examiners operating today are reliant upon a single commercial vendor to supply the bulk of their examination capabilities. They rely on one vendor for their core forensic platform and may have a handful of other commercial tools used for specific tasks that their main tool does not perform (or does not perform well). These experienced examiners who have little or no experience with open source tools will also hopefully benefit greatly from the content of this book.

## LAYOUT OF THE BOOK

Beyond the introductory chapter that follows, the rest of this book is divided up into eight chapters and one Appendix.

**Chapter 2** discusses the **Open Source Examination Platform**. We walk through all the prerequisites required to start compiling source code into executable code, install interpreters, and ensure we have a proper environment to build software on Ubuntu and Windows. We also install a Linux emulation environment on Windows along with some additional packages to bring Windows closer to "feature parity" with Linux for our purposes.

**Chapter 3** details **Disk and File System Analysis** using the Sleuth Kit. The Sleuth Kit is the premier open source file system forensic analysis framework. We explain use of the Sleuth Kit and the fundamentals of media analysis, disk and partition structures, and file system concepts. We also review additional core digital forensics topics such as hashing and the creation of forensic images.

**Chapter 4** begins our operating system-specific examination chapters with **Windows Systems and Artifacts**. We cover analysis of FAT and NTFS file systems, including internal structures of the NTFS Master File Table, extraction and analysis of Registry hives, event logs, and other Windows-specific artifacts. Finally, because

malware-related intrusion cases are becoming more and more prevalent, we discuss some of the artifacts that can be retrieved from Windows executable files.

We continue on to **Chapter 5**, **Linux Systems and Artifacts**, where we discuss analysis of the most common Linux file systems (Ext2 and 3) and identification, extraction, and analysis of artifacts found on Linux servers and desktops. System level artifacts include items involved in the Linux boot process, service control scripts, and user account management. User-generated artifacts include Linux graphical user environment traces indicating recently opened files, mounted volumes, and more.

**Chapter 6** is the final operating system-specific chapter, in which we examine **Mac OS X Systems and Artifacts**. We examine the HFS+ file system using the Sleuth Kit as well as an HFS-specific tool, HFSXplorer. We also analyze the Property List files that make up the bulk of OS X configuration information and user artifacts.

**Chapter 7** reviews **Internet Artifacts**. Internet Explorer, Mozilla Firefox, Apple Safari, and Google Chrome artifacts are processed and analyzed, along with Outlook, Maildir, and mbox formatted local mail.

**Chapter 8** is all about **File Analysis**. This chapter covers the analysis of files that aren't necessarily bound to a single system or operating system—documents, graphics files, videos, and more. Analysis of these types of files can be a big part of any investigation, and as these files move frequently between systems, many have the chance to carry traces of their source system with them. In addition, many of these file formats contain embedded information that can persist beyond the destruction of the file system or any other malicious tampering this side of wiping.

**Chapter 9** covers a range of topics under the themes of **Automating Analysis and Extending Capabilities**. We discuss the PyFLAG and DFF graphical investigation environments. We also review the *fiwalk* library designed to take the pain out of automated forensic data extraction. Additionally, we discuss the generation and analysis of timelines, along with some alternative ways to think about temporal analysis during an examination.

The **Appendix** discusses some non-open source tools that fill some niches not yet covered by open source tools. These tools are all available free of charge, but are not provided as open source software, and as such did not fit directly into the main content of the book. That said, the authors find these tools incredibly valuable and would be remiss in not including some discussion of them.

## WHAT IS NOT COVERED

While it is our goal to provide a book suitable for novice-to-intermediate examiners, if you do not have any experience with Linux at the command line, you may find it difficult to follow along with the tool use examples. While very few of the tools covered are Linux specific, most of the tool installation and subsequent usage examples are performed from a Linux console.

We focus exclusively on dead drive forensic analysis—media and images of systems that are offline. Collection and analysis of volatile data from running systems are not covered. Outside of the Linux platform, current tools for performing these tasks are largely closed source. That said, much of the analysis we go through is equally applicable to artifacts and items recovered from live systems.

Low-level detail of file system internals is intentionally omitted as this material is covered quite well in existing works. Likewise the development of open source tools is not discussed at length here. This is a book that first and foremost is concerned with the operational use of existing tools by forensic practitioners.

Outside of the Appendix, no commercial, proprietary, closed source, or otherwise restricted software is used.