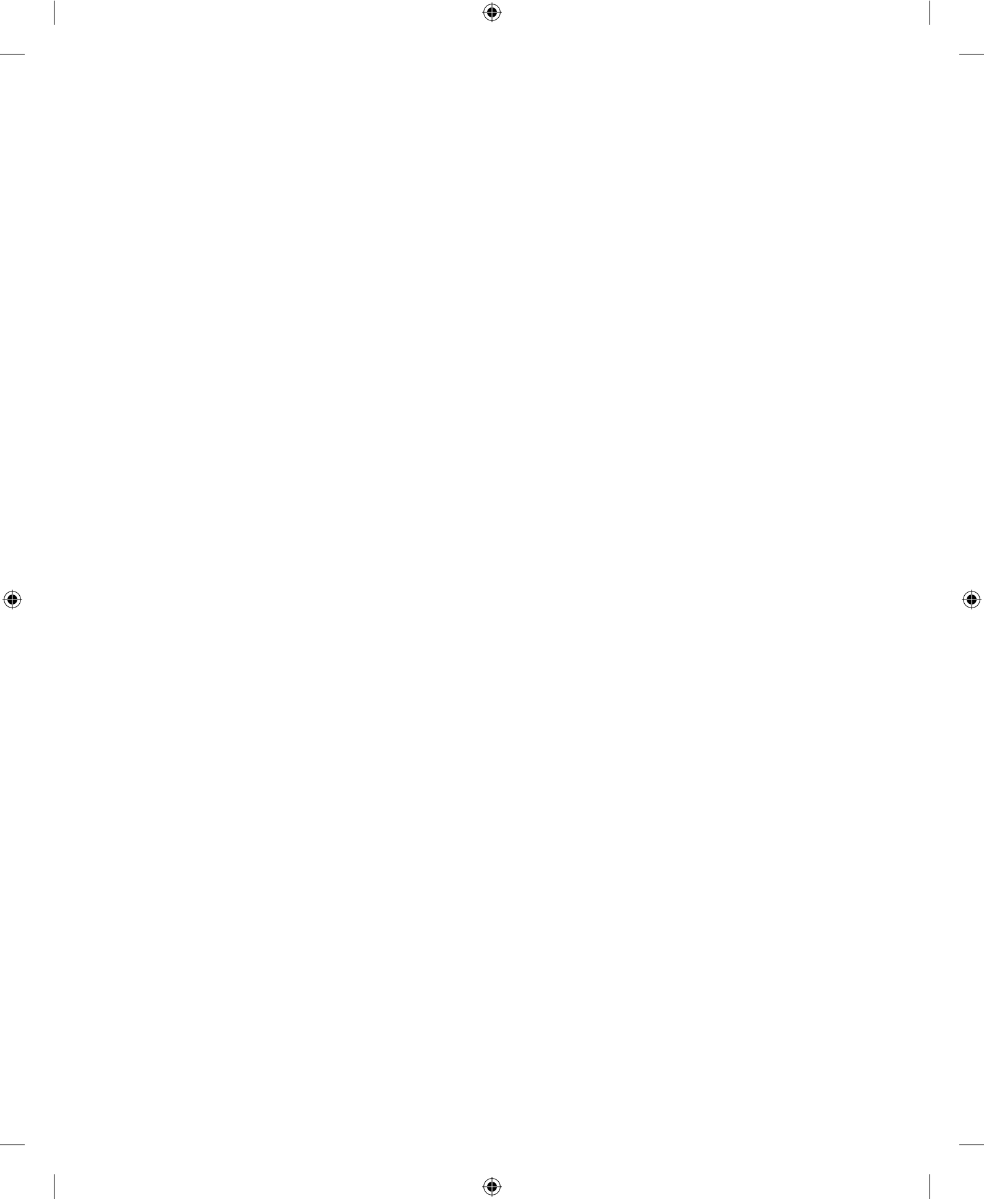


insert hyphen

# Data Driven **Security**



insert hyphen

# Data Driven **Security**

Analysis, Visualization,  
and Dashboards

comma / ?

JAY JACOBS  
BOB RUDIS

do:  
Jay Jacobs & Bob Rudis

WILEY



hyphen /  
comma / ?

Data Driven Security: Analysis, Visualization and Dashboards

Published by  
John Wiley & Sons, Inc.  
10475 Crosspoint Boulevard  
Indianapolis, IN 46256  
[www.wiley.com](http://www.wiley.com)

Copyright © 2014 by John Wiley & Sons, Inc., Indianapolis, Indiana

Published by John Wiley & Sons, Inc., Indianapolis, Indiana

Published simultaneously in Canada

ISBN: 978-1-118-79372-5

ISBN: 978-1-118-79366-4 (ebk)

ISBN: 9789-1-118-79382-4 (ebk)

Manufactured in the United States of America

10987654321

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 Sections 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, 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 646-8600. 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:** The publisher and the author make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation warranties of fitness for a particular purpose. No warranty may be created or extended by sales or promotional materials. The advice and strategies contained herein may not be suitable for every situation. This work is sold with the understanding that the publisher is not engaged in rendering legal, accounting, or other professional services. If professional assistance is required, the services of a competent professional person should be sought. Neither the publisher nor the author shall be liable for damages arising herefrom. The fact that an organization or Web site is referred to in this work as a citation and/or a potential source of further information does not mean that the author or the publisher endorses the information the organization or website may provide or recommendations it may make. Further, readers should be aware that Internet websites listed in this work may have changed or disappeared between when this work was written and when it is read.

For general information on our other products and services please contact our Customer Care Department within the United States at (877) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley publishes in a variety of print and electronic formats and by print-on-demand. Some material included with standard print versions of this book may not be included in e-books or in print-on-demand. If this book refers to media such as a CD or DVD that is not included in the version you purchased, you may download this material at <http://booksupport.wiley.com>. For more information about Wiley products, visit [www.wiley.com](http://www.wiley.com).

**Library of Congress Control Number:** XXXXXXXXXX

**Trademarks:** Wiley and the Wiley logo are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates, in the United States and other countries, and may not be used without written permission. ~~Insert third-party trademark information.~~ All other trademarks are the property of their respective owners. John Wiley & Sons, Inc. is not associated with any product or vendor mentioned in this book.

please  
provide



# About the Authors

(everyone else mentions twitter, please add):

“Jay can be found on twitter as @jayjacobs.”

**Jay Jacobs** has over 15 years of experience within IT and information security with a focus on cryptography, risk, and data analysis. As a Senior Data Analyst on the Verizon RISK team, he is a co-author on their annual Data Breach Investigation Report and spends much of his time analyzing and visualizing security-related data. Jay is a co-founder of the Society of Information Risk Analysts and currently serves on the organization's board of directors. He is an active blogger, a frequent speaker, a co-host on the *Risk Science* podcast and was co-chair of the 2014 Metricon security metrics/analytics conference. He holds a bachelor's degree in technology and management from Concordia University in Saint Paul, Minnesota, and a graduate certificate in Applied Statistics from Penn State.

**Bob Rudis** has over 20 years of experience using data to help defend global Fortune 100 companies. As Director of Enterprise Information Security & IT Risk Management at Liberty Mutual, he oversees their partnership with the regional, multi-sector Advanced Cyber Security Center on large scale security analytics initiatives. Bob is a serial tweeter (@hrbrmstr), avid blogger (rud.is), author, speaker, and regular contributor to the open source community (github.com/hrbrmstr). He currently serves on the board of directors for the Society of Information Risk Analysts (SIRA), is on the editorial board of the SANS Securing The Human program, and was co-chair of the 2014 Metricon security metrics/analytics conference. He holds a bachelor's degree in computer science from the University of Scranton.

# About the Technical Editor

**Russell Thomas** is a Security Data Scientist at Zions Bancorporation and a PhD candidate in Computational Social Science at George Mason University. He has over 30 years of computer industry experience in technical, management, and consulting roles. Mr. Thomas is a long-time community member of Securitymetrics.org and a founding member of the Society of Information Risk Analysts (SIRA). He blogs at <http://exploringpossibilityspace.blogspot.com/> and is @MrMeritology on Twitter.

rebreak



# Credits

**Executive Editor**

Carol Long

**Senior Project Editor**

Kevin Kent

**Technical Editor**

Russell Thomas

**Senior Production Editor**

Kathleen Wisor

**Copy Editor**

Kezia Endlsey

**Editorial Manager**

Mary Beth Wakefield

**Freelancer Editorial Manager**

Rosemarie Graham

**Associate Director of Marketing**

David Mayhew

**Marketing Manager**

Ashley Zurcher

**Business Manager**

Amy Knies

**Vice President and Executive Group Publisher**

Richard Swadley

**Associate Publisher**

Jim Minatel

**Project Coordinator, Cover**

Katie Crocker

**Proofreader**

Nancy Carrasco

**Indexer**

[Name]

**Cover Image**

[Name]

**Cover Designer**

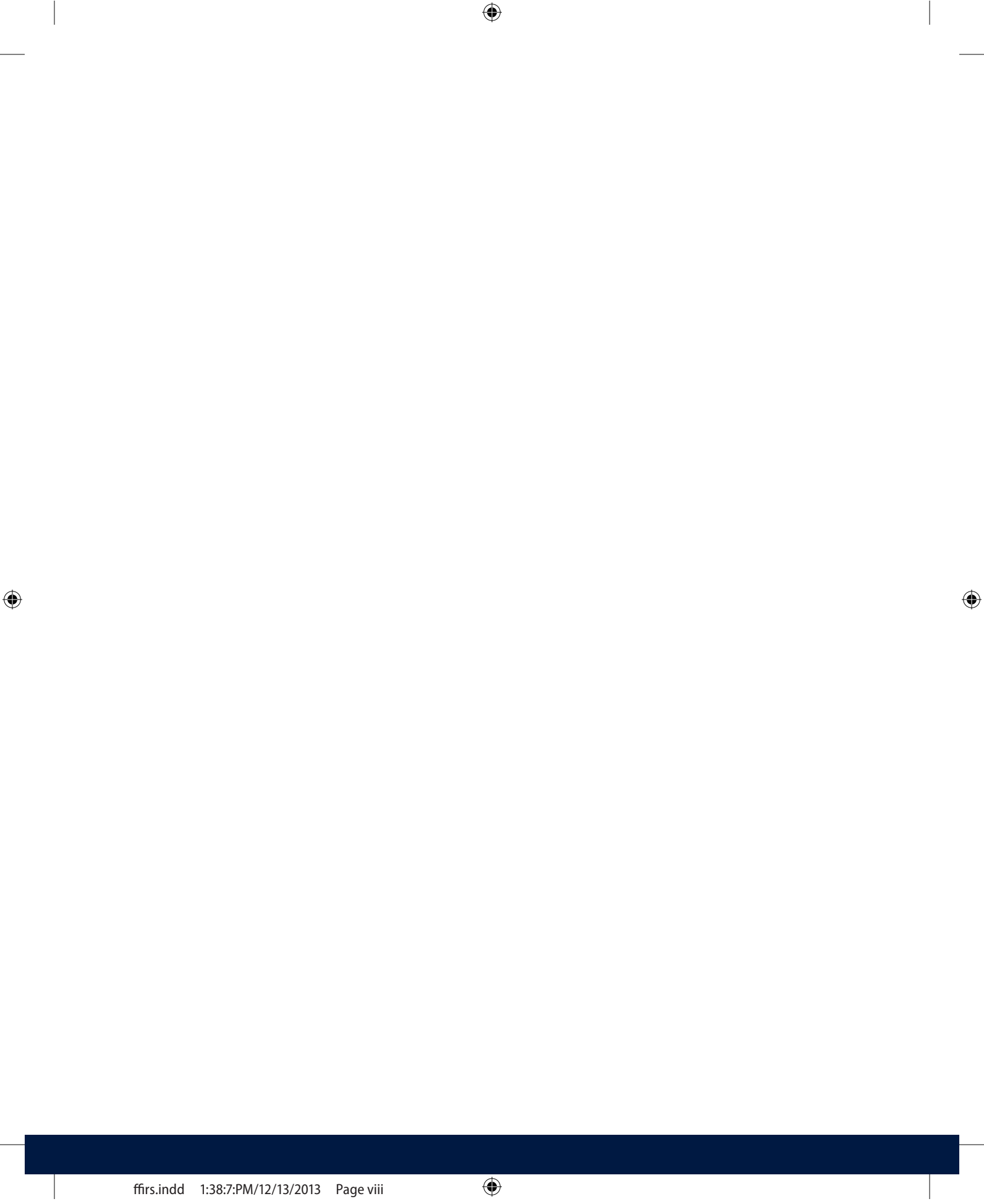
Ryan Sneed

Endsley / ?

Johnna VanHoose Dinse

Au: This was your image, correct?

Bob Rudis did the cover image





# Acknowledgments

While our names are on the cover, this book represents a good deal of work by a good number of (good) people. A huge thank you goes out to Russell Thomas, our technical editor. His meticulous attention to detail has not only made this book better, but it's also saved us from a few embarrassing mistakes. Thank you for those of you who have taken the time to prepare and share data for this project: Symantec, AlienVault, Stephen Patton, and David Severski. Thank you to Wade Baker for his contagious passion, Chris Porter for his contacts, and the RISK team at Verizon for their work and contribution of VERIS to the community. Thank you to the good folks at Wiley—especially Carol Long, Kevin Kent, and Kezia Endsley—who helped shape this work and kept us on track and motivated.

Thank you also to the many people who have contributed by responding to our emails, talking over ideas, and providing your feedback. Finally, thanks to the many vibrant and active communities around R, Python, data visualizations, and information security; hopefully, we can continue to blur the lines between those communities.

## Jay Jacobs

First and foremost, I would like to thank my parents. My father gave me his passion for learning and the confidence to try everything. My mother gave me her unwavering support, even when I was busy discovering which paths not to take. Thank you for providing a good environment to grow and learn. I would also like to thank my wife, Alicia. She is my best friend, loudest critic, and biggest fan. This work would not be possible without her love, support, and encouragement. And finally, I wish to thank my children for their patience: I'm ready for that game now.

## Bob Rudis

This book would not have been possible without the love, support, and nigh-unending patience through many a lost weekend of my truly amazing wife, Mary, and our three still-at-home children, Victoria, Jarrod, and Ian.

Thank you to Alexandre Pinto, Thomas Nudd, and Bill Pelletier for well-timed (though you probably didn't know it) messages of encouragement and inspiration. A special thank you to the open source community and reproducible research and open data movements who are behind most of the tools and practices in this text. Thank you, as well, to Josh Corman who came up with the spiffy title for the tome.

And, a final thank you—in recipe form—to those that requested one with the book:

### Pan Fried Gnocchi with Basil Pesto

- 2 C fresh Marseille basil
- 1/2 C fresh grated Romano cheese
- 1/2 C + 2 tbsp extra virgin olive oil
- 1/4 C pine nuts

check spelling

Make these  
H2s

“Ally” not  
“Alicia”

- 4 garlic scapes
- Himalayan sea salt; cracked pepper
- 1 lb. gnocchi (fresh or pre-made/vacuum sealed; gnocchi should be slightly dried if fresh)

Pulse (add in order): nuts, scapes, basil, cheese. Stream in 1/2 cup of olive oil, pulsing and scraping as needed until creamy, adding salt and pepper to taste. Set aside.

Heat a heavy-bottomed pan over medium-high heat; add remaining olive oil. When hot, add gnocchi, but don't crowd the pan or go above one layer. Let brown and crisp on one side for 3–4 minutes then flip and do the same on the other side for 2–3 minutes. Remove gnocchi from pan, toss with pesto, drizzle with saba and serve. Makes enough for 3–4 people.

# Contents at a Glance

align with rest of chapters

hyphen

align w/ title

Chapter 7:  
Learning From  
Security Breaches

Titles differ from tracking grid.

Chapter 9:  
Demystifying  
Machine Learning

Indroduction.....	xvii
Chapter 1 • The Journey to Data Driven Security .....	1
Chapter 2 • Building Your Analytics Toolbox: A Primer on Using R and Python for Security Analysis .....	21
Chapter 3 • Learning the “Hello World” of Security Data Analysis.....	39
Chapter 4 • Performing Exploratory Security Data Analysis .....	71
Chapter 5 • From Maps to Regression .....	103
Chapter 6 • Visualizing Security Data.....	137
Chapter 7 • Learning from Security Failures.....	161
Chapter 8 • Breaking Up With Your Relational Database .....	191
Chapter 9 • Machine Learning .....	217
Chapter 10 • Designing Effective Security Dashboards.....	245
Chapter 11 • Building Interactive Security Visualizations .....	269
Chapter 12 • Moving Toward Data-Driven Security .....	297
Appendix A • Resources and Tools .....	309
Appendix B • References .....	313
Index .....	321

This page should have bottom bar as in other pages?

