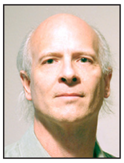
Interview with Dave Dittrich

RIK FARROW

Dave Dittrich is one of those rare people who started college declaring a major in photography and graphic arts but left with a computer science degree and went on to be involved in many "first in the world" cyber events. His incident response experiences, often involving personally identifiable information of both innocents and suspected computer criminals, led him to research the ethical and legal bounds within which "white hat" actors can justifiably act to respond to "black hat" hackers and criminals. He has written extensively on ethics and the "Active Response Continuum," served for six years on a University of Washington Institutional Review Board, and has recently been distilling this all into curricular resources for teaching practical ethical analysis. dave.dittrich@gmail.com



Rik is the editor of *;login:*. rik@usenix.org

I first met Dave Dittrich at USENIX Security in 2000. Dave had been working at University of Washington for many years by then and had made a name for himself with his analysis of malware installed on Internet-connected systems at the university.

I had learned about his work on distributed hacking tools, particularly the ones for carrying out distributed denial of service (DDoS) attacks. Someone within the NSA had kindly pointed me in that direction, and I had fortunately realized the potential impact and managed to get an article published days before MafiaBoy set off his big attack.

Rik Farrow: When did you start working in DFIR (Digital Forensics and Incident Response) at the University of Washington, and what was that like?

Dave Dittrich: My start in security came from the system administration side, out of necessity.

After working for a couple of years in the UW Chemistry Department, I took a position as the frontline UNIX workstation support contact for faculty and staff on campus. At the time, I think there was something like 20,000 UNIX workstations and maybe 3–4 times more Windows systems. But Windows didn't have a standard TCP/IP stack, so if a computer was broken into over the Internet, it would be a UNIX system. There were BSD, SunOS 3 and 4, System V, HP/UX, Irix, Digital UNIX, NeXT, and nascent Linux (Red Hat and Debian, mostly). I had to support them all, being the first (and usually only) person that would interface with the faculty and staff, relying on the University Computing Services system administrators and engineers for their experience when I didn't have it.

There would sometimes be dozens or hundreds of compromised systems at any given time, and I tried to help everyone as efficiently as possible. I took everything I learned and put it on my web page, and added it to the two-day R870 system administration course that I inherited from someone who retired right after I came on board [1]. I got bit-image copies of any interesting computer intrusion and got really efficient at forensic analysis using open source tools like Coroner's Toolkit by Dan Farmer and Wietse Venema, following public guidelines by the FBI and DoD, and developing my own investigative and reporting techniques.

RF: I think that UNIX `strings` was one of my favorite tools for a quick look at a suspicious binary. Coroner's Toolkit was amazing.

DD: Yeah, amazingly just using `strings` would be enough to get a pretty good idea from internal prompts, error messages, and system call identifiers of what a simple piece of malware was supposed to do. A disassembly could then provide some more detail. For example, is it a sniffer? A remote access trojan? A rootkit concealment tool? An exploit?

Another really basic technique, but one that I don't see commonly used by forensic analysts, is using file system Modify/Access/Create (MAC) timelines to develop situational awareness about post-intrusion activity. Forensic analysts often search for "known bads" using hash databases, or exclude programs based on "known goods" hash databases, or search for known Windows Registry keys, etc. In other words, looking for things based on simple signatures (often signatures derived by others at different sites). While this might work, it also might take hours of indexing to come up with nothing, especially if the malware is polymorphic or crafted specifically for that victim, meaning nobody else would see the same binary in their generalized threat intelligence telemetry. Or it might find several artifacts from different unrelated intrusions over time, confusing the analyst. Just finding a hash match or a file name match is the start of an analytic process, not the end.

It is really hard to effectively wipe out all possible evidence of compromise of the integrity of a computer system. I think it's safe to say that most intrusions up to the early 2000s had almost no effort spent on advanced concealment and wiping of fingerprints, so to speak. Rootkits were very common (both user level and kernel level) but were usually pretty easy to defeat if you know how the operating system, file system, and network connections behave. But you need to be able to show your work and prove it to a "preponderance of the evidence" in civil cases, and "beyond a reasonable doubt" in a criminal case.

I have found it far quicker and more useful to leverage initial facts (including time and date of suspected malicious activity) to find the directory where malware was initially dropped or where configuration files and/or log files are stored. In situations where there is no enterprise endpoint protection agent in place, a very common situation, you need to "live off the land" in terms of evidence collection. To increase confidence, you then include external sources of evidence to confirm/refute things like clock skew, missing the year or time zone in system log lines, etc.

I developed a forensic analysis and reporting methodology using the tools and techniques described by Farmer and Venema in the notes from their 1999 IBM forensic training event. I described this technique and how to use it in a guide I published later that year, "Basic Steps in Forensic Analysis of UNIX Systems" [2]. I also used this technique in a two-hour "house call" on the University of Washington campus network to quickly get around a kernel-level rootkit on a Linux server, which became the chapter "Omerta" in Mike Schiffman's book *Hacker's Challenge* [3].

The owner of that system was 100% sure his system was not compromised, since the kernel-level rootkit worked so well. I hooked his computer and my laptop up to a hub and showed him the IRC bot traffic coming from a process that wasn't listed in `netstat`, or `ps` output. I then had him run `dd` using `netcat` to pipe the root partition to my laptop, where I used the Coroner's Toolkit to get a MAC timeline and later to extract and analyze deleted file space. Having obtained a bit-image copy of the root partition to preserve any evidence, it only took a short time, while simultaneously copying the other partitions, to identify and disable the rootkit. All of the malicious processes now showed up!

By the next day I had a full understanding of what had happened, identified all of the other systems around the world being used by the group from network traffic and internal log or rootkit configuration files, and reported to all the other victim sites and to CERT/CC.

Farmer and Venema, two of the voices of reason in the forensic arena, published a much more detailed description of the underlying operating system behaviors and file-system functions that preserve evidence in their book, *Forensic Discovery* [4]. They showed in technical detail how, despite file deletions (or even re-installation of the operating system, if you look hard enough!), you can do the same kind of analysis that geologists do to understand the history of a specific location by examining the composition of soil layers, rock or shell inclusions, discontinuities in soil layers, etc. With an understanding of how kernels running programs affect MAC times in each type of file system in use, you can not only make quantifiable conclusions based on interpretation of MAC timelines, but you can demonstrate through experiments using the same kernel and file system that you can reproduce the results to show proof to back up your theory!

If your objective is to support criminal process, this is very important in order to meet an evidentiary standard known as the "Daubert Standard" (Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579 (1993)): Federal Rules of Evidence 702 requires that an expert witness should possess the kind of knowledge as found in Farmer and Venema's book, use that knowledge to help the court understand the evidence or determine facts at issue, base their testimony on sufficient facts or data that were the product of reliable principles and methods, and reliably apply those principles and methods to the facts of the case.

RF: Another thing I recall you were involved with was the Honeynet Project (HP). I asked Lance Spitzner to write about the HP in 2002 [5]. How did you get involved in the HP?

DD: My publications and conference talks on sniffers, rootkits, post-intrusion log alteration and concealment, and DDoS handler/agent tools, pre-cursors to today's "botnets," got me an invitation to the Honeynet Project. The first publications we did referenced many of the whitepapers I published on my UW home page.

People kept saying, "It's great that you mention how to use tools and how to analyze compromised systems, but I don't have a honeypot set up and want a bit image disk copy to work with." So Lance Spitzner asked me to organize the Forensic Challenge so that people could have a real-life compromised Linux system to work with. I spent over a hundred hours in one month doing the reference analysis, setting up the rules, organizing the judges, and managing the judging process. It was the top most-popular download on the HP web site for a few years running!

I also organized the Reverse Challenge, which turned out to be yet-another-DDoS-bot!

RF: How did you get involved with working on the Menlo Report? Is that something you and Erin Kenneally decided to do on your own?

DD: Erin and I both came into our roles in the process from previous DHS and ethics (and legal, in Erin's case) work we had done.

I had been working within the PREDICT project (a research data repository project, now known as the Information Marketplace for Policy and Analysis of Cyber-risk and Trust, or IMPACT) at DHS for many years. Around 2006, I was trying to develop honeypot images and related logs and network traffic for use in research. This kind of sandbox processing of malware artifacts is commonplace today, but not in the mid-2000s. One of the larger such botnet-related dataset collections today is maintained by the Czech Technical University in Prague (https://mcfp.felk.cvut.cz/publicDatasets/).

One of the botnets I had studied, known as "Nugache," was written up in USENIX *;login:* in 2007 [6]. Nugache had some features far in advance of the most visible botnet in the world at the time, the "Storm botnet." I was keeping a close eye on Storm and the differences in Nugache that really had me worried due to the level of apparent sophistication in that botnet (that wouldn't be publicly shown to have been surpassed until the Conficker.C variant came out years later).

I saw the December 2008 CCC presentation "Stormfucker: Owning the Storm Botnet" by researchers from the University of Bonn, inspired by research from the University of Mannheim, where they demonstrated a partially tested implementation of software components necessary for constructing a "white worm" that could be released to clean up Storm botnet-infected nodes. Afterwards, I began writing on the ethics of cleaning up botnets. This followed on the Active Response Continuum research I had done, and my take on the ethics was very applied and focused on the overlap of research and operations (including law enforcement investigations), not just a pure academic research perspective.

My first attempt at publication at the USENIX LEET '09 workshop was rejected, but I was invited to participate on a panel entitled "Ethics in Botnet Research" in April 2009 [7].

That initial rejected paper grew and became a technical report co-authored with Michael Bailey (a PREDICT Principal Investigator) and Sven Dietrich (whom I had been working with on Nugache). We released the technical report the same day as the LEET panel [8].

I had several people I knew with ethical review experience review the paper and case studies to see if they would even require research ethical review. One was Katherine Carpenter, with whom I've subsequently written several articles and papers. The other two were Tanya Matthews and Shannon Sewards, who worked at the University of Washington's Institutional Review Board (IRB). I also joined one of UW's IRB committees to learn how the process works from firsthand experience, serving on the committee for over six years.

Doug Maughan was at the LEET panel and invited me to speak about this paper at the first workshop on ethics in ICT research he was setting up for the next month (May 26-27, 2009). That workshop led to formation of the Menlo Working Group. The technical report I co-authored with Bailey and Dietrich was provided to the Working Group and served as some of the background and case studies for the Companion to the Menlo Report.

RF: I believe you wrote about the process. It's enough to say that many people were involved, but you and Erin created the report that got published in the Federal Record.

DD: The process was covered in an *IEEE Security and Privacy* article [9].

We had a large Working Group, approximately two dozen people, a similarly sized group of external reviewers, and a number of official responses to the publication in the Federal Register that had to be integrated and summarized in an official response in the Federal Register. I learned a lot about the Federal Register and its relationship to federal regulations!

Erin Kenneally was serving as legal counsel to CAIDA (another PREDICT performer). Erin and I both had the capacity to wrangle the report drafting and commenting/editing process. We got closer to a final draft ready for submission to the Federal Register and subsequent public response when Michael Bailey joined us to help out with the final push (and to work with us to start publishing and speaking about the Menlo Report as part of the outreach process).

RF: What else were you doing at UW?

DD: A couple years after that, Mike Eisenberg (Dean of the Information School) and David Notkin (chair of the Computer Science and Engineering Department) bought out half of my time to allow me to reach out to other universities and community colleges, get the UW accepted into the National Security Agency's Center of Academic Excellence in Information Assurance Education (CAE IAE) program, help start up the Center for Information Assurance and Cybersecurity (CIAC), and begin a career as a staff research scientist with permission to be a Principal Investigator on grants, despite only having a BS degree. I owe a great deal to Mike Eisenberg.

Over the next 10+ years, I brought in over $4 million in grants and contracts and covered my salary and that of a few others at various times. My first grant was from Cisco Systems Critical Infrastructure Assurance Group to study *active defense*. I coined the term *Active Response Continuum* (ARC) to make it clear this is not a black/white situation by any means but, rather, a set of ranges or levels (capacity to respond, aggressiveness, intrusiveness, risk, etc.). I collaborated with Kenneth Einar Himma to write one of the early papers on the topic (http://ssrn.com/abstract=790585) and presented first publicly at AusCERT 2005.

We came at the subject from the perspective of private-sector response and framed it in terms of ethical principles, as opposed to the military *law of war* context taken by most publications on the topic to date. The concept of ethics in security operations and research has remained a central part of my research and publications since then. The bulk of the funding I secured at the UW was from Doug Maughan (another person to whom I owe a great deal) at DHS but also included grants or contracts from NSF, DoD, the FTC, and industry.

Over that same period I had permission to work on outside contracts and pro bono projects, including contract support to criminal defense lawyers, federal public defenders, assisting a few DDoS victims, assisting the Federal Trade Commission on a fake-drug civil temporary restraining order (TRO) case, and providing declarations to the court in two of Microsoft's ex-parte TROs in the Waledac and Rustock botnet cases.

RF: What are your plans for the future?

DD: I'll be really honest: I'm figuring that out. Let me explain.

During my last major project as a Principal Investigator at the UW, I worked so hard I was burning myself out. Physically, I have a nerve impingement in my neck that began to cause pain, tingling, and numbness in my back, shoulder, and arm. Emotionally, I was taking on too much stress (which combined with the physical issues to produce a negative feedback loop). My doctor, friends and family were all telling me I had to cut back, change my work habits, and take it easier to begin to recover.

I just read Arthur C. Brooks' *Atlantic* piece, "Your Professional Decline Is Coming (Much) Sooner Than You Think: Here's How to Make the Most of It" [10]. His article really spoke to me and made me realize some things that have been in the back of my mind lately.

I've recently been writing a history of the early days of the Honeynet Project, going back over some of the things I did in the late 1990s and early 2000s. This August 19th is the 20th anniversary of the first massive DDoS (handler/agent style) attack on the University of Minnesota that lead me to write the first DDoS tool analyses. It surprised me a little to realize just how much I did in those days (all the DDoS tool analyses I wrote, computer security incidents I investigated and reported on to CERT/CC and the FBI, projects taking up hundreds of hours over a month or two, trips and talks and publications, all on top of a 40+ hour work week). As Brooks describes, I made my name and reputation in this industry using fluid intelligence and a dedication to serving the public through open source research, digital forensics, malware analysis and threat intelligence, and publication. But I am learning (the hard way) how unsustainable that level of productivity really is. I feel confident I can still identify and solve novel "cyber" problems, but physically I can't put in 12+ hour days any longer.

Over the last two years I've started shifting to, as Brooks puts it, using crystallized intelligence—applying all the lessons I've learned in information security over the decades, the recommendations and predictions I've made, all that I have read and researched, the linkages I'm capable of recognizing—as a contract subject matter expert and an author. I have invested thousands of hours in producing open source tools, documenting ways to solve some basic information security problems that have persisted for decades (like default passwords and secrets leaked through source code repositories), and combining case studies and other material from papers I've written and the Menlo Report effort to produce materials for a full-day applied ethics tutorial/course. Ever since my UNIX workstation support days, I have tried to teach what I have learned by including what-to-do and how-to-do-it information in my publications and have given many talks and guest lectures. So perhaps teaching is my future? After all, my father (before he passed away) and my older brother today both taught college physics as professors for decades each.

I'm excited for the next direction my professional and personal life will take, similar to the way I used to feel in my 30s and 40s when preparing to go on a multi-day backcountry ski trip. I know how to navigate finding a route, the general direction I want to go, and I've done all the preparation and accumulated the requisite knowledge. But I don't know right now precisely what path I will take, what kind of objective hazards I will have to overcome, the amazing views from the summits, or what pleasures (and discomforts) I will encounter on the way.

References

[1] "R870: UNIX System Administration—A Survival Course": https://www.washington.edu/R870/cover-page.html.

[2] D. Dittrich, "Basic Steps in Forensic Analysis of UNIX Systems": https://staff.washington.edu/dittrich/misc/forensics/.

[3] M. Schiffman, *Hacker's Challenge* (McGraw-Hill, 2002).

[4] D. Farmer, W. Venema, *Forensic Discovery* (Addison-Wesley Professional, 2005).

[5] L. Spitzner, "HOSUS (Honeypot Surveillance System)," *;login:*, vol. 27, no. 6 (USENIX, December 2002): https://www.usenix.org/system/files/login/articles/1252-spitzner.pdf.

[6] S. Stover, D. Dittrich, J. Hernandez, and S. Dietrich, "Analysis of the Storm and Nugache Trojans: P2P Is Here," *;login:,* vol. 32, no. 6 (USENIX, December 2007): https://www.usenix.org/system/files/login/articles/526-stover.pdf.

[7] J. Brodkin, "The Legal Risks of Ethical Hacking," *Network World,* April 24, 2009: https://www.networkworld.com/article/2268198/the-legal-risks-of-ethical-hacking.html.

[8] D. Dittrich, M. Bailey, S. Dietrich, "Towards Community Standards for Ethical Behavior in Computer Security Research," Technical Report CS 2009-01, Stevens Institute of Technology, April 2009: https://staff.washington.edu/dittrich/papers/dbd2009tr1/dbd2009tr1.pdf.

[9] M. Bailey, D. Dittrich, E. Kenneally, and D. Maughan, "The Menlo Report," *Security and Privacy*, vol. 10, no. 2 (IEEE, March/April 2012), pp. 71–75: http://www.caida.org/publications/papers/2012/menlo_report/menlo_report.pdf.

[10] A. Brooks, "Your Professional Decline Is Coming (Much) Sooner Than You Think: Here's How to Make the Most of It," *The Atlantic*, July 2019: https://www.theatlantic.com/magazine/archive/2019/07/work-peak-professional-decline/590650/.

Save the Dates!

usenix
THE ADVANCED
COMPUTING SYSTEMS
ASSOCIATION

29TH USENIX SECURITY SYMPOSIUM

August 12–14, 2020 • Boston, MA, USA

The 29th USENIX Security Symposium brings together researchers, practitioners, system administrators, system programmers, and others to share and explore the latest advances in the security and privacy of computer systems and networks.

The Symposium will span three days, with a technical program including refereed papers, invited talks, posters, panel discussions, and Birds-of-a-Feather sessions. Co-located workshops will precede the Symposium on August 10 and 11.

Paper submission deadlines:
Fall Quarter: Friday, November 15, 2019
Winter Quarter: Saturday, February 15, 2020

Invited talk and panel proposals deadline:
Friday, February 14, 2020

Program Co-Chairs

Srdjan Capkun
ETH Zurich

Franziska Roesner
University of Washington

www.usenix.org/sec20

Sixteenth Symposium on Usable Privacy and Security

Co-located with USENIX Security ’20
August 9–11, 2020 • Boston, MA, USA

The Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020) will bring together an interdisciplinary group of researchers and practitioners in human computer interaction, security, and privacy. The program will feature technical papers, including replication papers and systematization of knowledge papers, workshops and tutorials, a poster session, and lightning talks.

Symposium Organizers

General Chair
Heather Richter Lipford,
University of North Carolina at Charlotte

Technical Papers Co-Chairs
Michelle Mazurek, *University of Maryland*
Joe Calandrino, *Federal Trade Commission*

www.usenix.org/soups2020