the public interest hack

How are hacking and leaking related? **Gabriella Coleman** introduces us to the "public interest hack" and explains how it emerged.

ILLUSTRATION: AMISHA GADANI

IN WINTER OF 2014, AN INQUIRY FROM A JOURNALIST LANDED IN MY email inbox. His message opened innocently: "how can I can safely and effectively get plugged into the hacker community?" He continued with a hypothetical explanation: "my sense is that folks in particular hacker circles might be interested to know that a reporter is digging on topic X because they too are eager to see a spotlight thrown there." Further discussion revealed he was interested in how hackers chose targets, a query prompted by the recent spate of Anonymous-led hacks and document leaks. After reminding him it was illegal to seek hacking aid of any sort, I told him that, as far as I knew, the hackers themselves had initiated these computer infiltrations and subsequent document exfiltrations. There was no indication they were ever prompted by a journalist or other citizen—as it should be, I stressed.

I was satisfied I had relayed to him (and any snoopers listening in) my unambiguous objections to such a scheme; but at the time, I overlooked the historical benchmark furnished by his inquiry. That he was interested in how hackers went about landing documents to publish signaled that a new strategy—what I am calling the public interest hack—had, by this historical juncture, become fully imaginable and established. To define it in its simplest terms: a public interest hack (PIH) entails a computer infiltration for the purpose of leaking documents that will have political consequence. Rather than perpetrating a hack just for hacking's sake, as hackers have always done, the PIH is a hack that will *interest the public* (even if, as we have seen with the DNC hacks and the Macron hack, it is not necessarily 'in the public interest' in some simply positive sense). This tactic can resemble traditional forms of leaking and whistleblowing, like the Pentagon Papers, insofar as both are high-risk activities leading to the release of publicly relevant documents. But they are distinct: because the PIH conjoins a computer intrusion—advertised as such—with a particular type of leak. The PIH has also taken on two distinct forms. In one class, many of the most prominent cases of the last five years—the hacks against security or intelligence firms like HBGary, Stratfor, Hacking Team, and FlexiSpy were orchestrated by hacktivists who explicitly sought to expose wrongdoing. Another class—like the Guardians of Peace hack of Sony Pictures and Guccifer 2.0's hack of the Democratic National Convention—were carried out by mysterious crews who, in contrast, have obscured their intentions but still released data and documents that spurred extraordinary public attention and inquiry (See table 1).

When I have told hackers or technology journalists of my hunch—that the PIH strategy did not really quite exist prior to 2007 and was largely indebted to Anonymous—none of them believed me. In fact, I did not believe it myself. It is why I tapped these experts in the first place seeking to find the esoteric or

overlooked cases when hackers had infiltrated a system, snatched documents, and published them widely, triggering substantial news or inquiry about the hack *and* the documents. But all they could offer, along with their skepticism, were many cases of hacktivist interventions ("how about the NASA Wank Worm?" they might charge back, among *many* other cases). None, however, quite fit the mold of this relatively new political strategy. What then distinguishes the PIH from other varieties of hacks, leaks, and whistleblowing? And why did it come into being only between 2007 and 2011, when it was conceivable—ideologically and materially—for it to have emerged much earlier?

Prior to the emergence of the PIH, various kinds of hack-leak combinations and hacktivist techniques were common. Indeed, most of these can qualify as both political and of interest to the public; but as proposed here, the PIH is a more distinctive and a more singular category that excludes the great majority of hacks, leaks and breaches. For instance, hackers have long infiltrated systems for all sorts of reasons—for fun, learning, and showmanship—and in the process swiped data and documents but never released them. For decades, hackers have also acquired and published credentials: passwords, log ins, and credit card numbers. But such leaked information can only be mobilized in the narrow form of consumer security advocacy. It is a distinct strategy for hackers and security researchers to use high profile breaches to urge corporate executives or government officials to invest more resources into digital fortifications. Other hacktivist techniques, like website defacements, distributed denial of service (DDoS) attacks, and even hacks of sabotage (like deleting files), don't entail document acquisition; thus they fail to qualify as a public interest hack. Finally, some black hat hackers were known to shame enemies by acquiring a victim's email spool and publishing it; but these events tended to be personal revenge skirmishes, with the emails never meant for wider uptake. Many examples

TABLE 1: Public Interest Leaks

TARGET	MATERIALS	HACKERS	YEAR
Hal Turner	limited emails leak	Unknown	2008
Sarah Palin	screen shots of emails	David Kernell	2008
Climatic Research Unit at University of West Anglia	emails	Unknown	2009
HBGary Federal	emails	Anonymous	2011
Stratfor Intelligence	emails	Anonymous	2011
Syrian Government	emails	RevoluSec	2012
Gamma Group	technical documentation and software	Phineas Phisher	2013
Hacking Team	emails	Phineas Phisher	2014
Peruvian government	emails	Anonymous/Lulzsec	2014
Sony Pictures	emails, documents, movies	Guardians of Peace	2014
CIA Director John Brennan	emails	Cracka's with Attitude	2015
Turkish AKP leaks	emails	Phineas Phisher and others	2016
Bradley Foundation hack	emails	Anonymous Globo	2016
Democratic National Convention	emails	Guccifer 2.0	2016
Colin Powell	emails	Guccifer 2.0	2016
John Podesta	emails	Guccifer 2.0	2016
Cellebrite	documents and circumvention tools	Unknown	2017
Retina and FlexiSpy	source code, HR documents, and other files	Decepticons	2017
Emmanuel Macron campaign staff	emails	Unknown	2017

of this class of hacks may well exist, but most have never pierced public consciousness.

There are likely cases—and I know of one—whereby a hacker leaked documents to the public or a journalist but did so without advertising the source of the material as a hack. I've identified one occurrence in this style, but had to venture deeply to recover it: in the mid-1990s, amid trenchant critiques of the Church of Scientology voiced on a popular Usenet mailing list, a hacker accessed Scientology servers, siphoned some documents and released them to the list. Still, measuring this early hacking case by my criteria, this hack-leak fails to qualify as a public interest hack because the hacker never advertised how he or she obtained the material. If we compare this instance with those hacks and leaks orchestrated by Phineas Phisher—who after hacking the Italian firm Hacking Team, published a "Hack Back" manual (2016) seeking to galvanize others to emulate him—we can identify the precise historical period when hackers publicized this strategy and thus positioned it for adoption and replication.

The PIH stabilized only in 2011, an exceptional year of political ferment characterized by waves of street-based demonstrations and the ascendancy of the hacker as a major geopolitical force. With Anonymous and WikiLeaks, hackers pushed the levers of power in new and far more consequential ways, making hacks and leaks the stuff of foreign policy briefs and international relations debates. In this period, Anonymous hackers twice stumbled upon newsworthy documents that they then published on accessible platforms like the Pirate Bay or WikiLeaks. Their conspicuous brand of hacking—accompanied by catchy digital posters and videos—lured in media professionals who boosted Anonymous' profile and by extension raised the profile of this mode of disclosure, ensuring that scattered instances of this method would crystallize into a template for emulation. But before we turn to Anonymous proper and the stabilization of this tactic, let's start with the pre-history of this method.

A BRIEF GENEALOGY OF THE PUBLIC INTEREST HACK

Roughly five years before hackers executed one of the first instances of the public interest hack, there are two borderline cases that prefigure the tactic: the acquisition of digital documents from the voting machine company Diebold; and the publication of emails from the now-defunct energy giant Enron. Even if these materials were not obtained by hacks, these two cases drew a hermeneutic circle, making it apparent that such digital information might be out there for the taking. The events also signaled that releasing digitally-hosted or digitally-compiled data, like emails, could potentially serve a democratic function by exposing or corroborating wrongdoing.

The Diebold case began in 2002 when Seattle resident Bev Harris learned that her county had purchased touchscreen voting machines and she flung herself into research on software vulnerabilities. While seeking technology experts online who could answer her litany of concerns, she found something more consequential: the source code, hardware schematics, internal mailing list archives, passwords, and documents for vote-counting software. Her initial attempt to hand over the documents to journalists by sending "more than 100,000 bulletins directly to the appropriate editors and producers," proved ineffectual (Harris 2003: 158). Later in the year, spurred by research enabled by the documents, the *New York Times* finally ran an exclusive story on "the stunning research flaws" in the Diebold system (Schwartz, 2003).

In same period, a large corpus of corporate emails—over 600,000 emails written by Enron employees—were published for the first time on the internet. The responsible party was not a reckless hacker, WikiLeaks (an organization not yet in existence), nor the Russian government, but an obscure American government agency: the Federal Energy Regulatory Commission (FERC). At the time, Enron had been embroiled in scandal with its top corporate executives under investigation for fraud. According to a later report by the *Wall Street Journal*, the FERC published the corpus "to help the public better understand whether Enron helped to create— and then profit from—an energy shortage in California during 2000 and 2001" (Berman, 2003). In spite of FERC's intentions, the contents of the emails attracted scant journalistic scrutiny; the *Wall Street Journal* rebuffed the release at the time, citing privacy violations. Not long after, another journalist defended the publication of the emails for offering a glimpse—into the "soul" as he put it—of the corrupt organizational culture of Enron (Grieve, 2003).

A few years later in 2007, a retributive attack orchestrated by anonymous 4chan users marked one of the first instances of a hack where information found in exfiltrated emails was publicized to damage the reputation of a targeted individual *and* picked up by organizations well outside of the technology and hacker community. It all began when Anonymous trolls prank called Hal Turner, a white supremacist radio host. When he made the grave error of doxing the callers, a group of 4chan anons decided to dox him right back: broadcasting Turner's home phone number, previous places of residence, and criminal records. As the doxing feud escalated, online allegations swirled that Turner was an FBI mole, sleuthing for the government to out white supremacists. The ostensible source of the accusation came from emails acquired by anonymous (not Anonymous) hackers. While the emails have since vanished, and didn't at the time spark any stories in the mainstream press, they became public knowledge, as groups like the Southern Poverty Law Center posted news of the hack and emails on their website (Potok, 2008). Due to this hack and leak, Hal Turner—once a beloved public personality among hard-core racists—became a pariah.

Another similar incident was executed in 2008 by David Kernell and was directed against then-presidential candidate Sarah Palin. Kernell, revealed to be the son of a Democratic representative, hacked Palin's Yahoo email account and posted to 4chan proof of the intrusion, an explanation of why and how he carried out the hack, and his fears of getting caught (his hunch proved founded: he was arrested not long after). Sharing a few screenshots he lamented, "there was *nothing* there, nothing incriminating, nothing that would derail her campaign as I hoped" (Schor, 2008). Even though he found nothing to publish, the case signals that by 2008, hackers were openly pursuing this game plan; and unlike the Hal Turner incident, the mainstream press picked up the Sarah Palin hack, with Gawker and WikiLeaks

republishing the screenshots. The wide coverage likely worked to sow the idea for future uptake.

Another two years would pass before Anonymous would strike again with a hack leading to a newsworthy email disclosure. In 2010, they managed to publish a large email cache thanks to a technical bungle by a company targeted by Anonymous hacktivists for other reasons. In September 2010, an Anonymous activist node by the name of AnonOps launched a pro-piracy campaign by hammering a slew of copyright industry websites with DDoS attacks. One target was ACS:Law, a British law firm under fire for sending thousands of notices to British citizens threatening them with lawsuits unless they ponied up a lump sum for their alleged piracy. ACS:Law's emails were obtained when, in the midst of being thrashed with a DDoS attack, ACS:Law removed their website from the internet. Upon restoring it, a misconfiguration meant all their email was on deck, available for the taking. Anonymous snapped up the digital assets and re-directed the emails to The Pirate Bay. Various parties waded into material, including Ars Technica reporter Nate Anderson, who provided an in-depth exposition of the emails, laying out the company's workings—and how many of their threats to supposed pirates were recklessly targeted (Anderson, 2010). "If there's one great theme running through these letters," highlighted Anderson, it's the poverty of the respondents" (ibid). Ultimately, the consequences of this hack and email disclosure were as direct as they were substantial: the government levied fines against the firm for its poor security and failure to protect sensitive personal data, and the firm was forced to close.

From the ACS:Law leak onwards, it became clear that the act of publishing exfiltrated digital content would garner public attention and—depending on the nature of the content—could serve particular political interests, in this case defending ordinary people from aggressive anti-piracy corporations. Hackers affiliated with Anonymous—and eventually others—at this moment became more deliberate: directing their finely-honed skills towards intelligence gathering of leakable information. For instance, in January 2011, some of the same hackers who published the ACS:Law emails squirreled into the Tunisian Prime Minister's email servers, hoping to find damning material that if released could turbo-boost the popular revolt gripping the nation. Their jaunt proved unsuccessful and they had to remain satisfied with the consolation prize of various website defacements.

That is, until two weeks later when the same platoon hacked Aaron Barr, the CEO of federal intelligence firm HBGary Federal. Barr was on a quest to infiltrate and dox Anonymous hackers. After the *Financial Times* published a piece detailing Barr's crackpot plan to publicly identify the core leaders driving the hacking operations (Menn, 2011), these hackers snarled back at Barr (whose 'intel' was wrong) with their own merciless brand of "infiltration." In one evening, Anonymous hackers snaked their way into HBGary Federal computer systems, hauled away the company's emails, posted them on The Pirate Bay, and gutted whatever else remained on the system.

HB Gary CEO Aaron Barr on The Colbert Report, 2011.

Owing in part to the irony of a ragtag band of hackers taking down a security firm with minimal effort, and the damning plot discovered in the emails, Operation HBGary became legendary among hackers and security professionals. The emails were full of fascinating information—including a PowerPoint concocted by Barr in partnership with Palantir and Berico employees, detailing plans to thwart and destroy WikiLeaks and its associates using dubious and illegal methods. One of the more reprehensible tidbits of their plan was to slander journalist Glenn Greenwald who, according to their assessment, would halt supporting WikiLeaks if his career was put under jeopardy.

Because the email contents and the logistics of the hack were juicy, shocking, and newsworthy—tailor-made for our contemporary media environment—the HBGary hack and leak dominated the news cycle for days. And like ASC:Law before it, airing the emails had an impact far beyond the shame it bestowed upon Aaron Barr. Disgraced, he was forced to resign; and not long after, HBGary Federal was itself dismantled. In the euphoria of victory, these hackers were emboldened to hack even more, which is precisely the path they took: first with a breakaway group Lulzsec and later with Antisec.

The gale of Anonymous hacking in 2011 brought seasoned hacktivist Jeremy Hammond out of retirement. Chartering a militant crew, Antisec, Hammond ensured that under his tutelage Anonymous would continue to prowl servers for the acquisition of incriminating evidence destined for wider distribution. After a string of hacks, one audacious exfiltration finally resulted in his arrest by the FBI. Rolled out against an intelligence firm Stratfor, the hack landed Strafor's emails, which Hammond sent to WikiLeaks. Journalists then mined them for evidence, pointing to the corporate spying against activists. Unlike the HBGary hacks, here Hammond and his teammates were not triggered by revenge—acting merely reactively—but were instead proactively seeking information.

Before the HBGary and the Stratfor hacks, hackers had certainly started to intrude systems for the purpose of extracting the sort of information the public or journalists might deem important. But the few successful instances of such an approach were scattershot or obscure. From this moment on in 2011, a time period when hacktivism itself had soared into the geopolitical stratosphere, this tactic gained momentum and seemed to settle into political pattern. The HBGary and Stratfor hacks were a sign a new threshold had been reached, at least in North American, European, and Latin American regions,[1] but it was not entirely clear whether the PIH would survive after law enforcement arrested scores of hackers who were responsible for these types of hacks.

The answer came in 2014 when other hacktivists executed exceptionally visible and high-impact public interest hacks. In Peru, the government nearly dissolved after a two-person Anonymous hacktivist crew, Lulzsec Peru,

distributed hacked emails from the Department of the Interior—correspondence teeming with evidence of corruption. After a flurry of press coverage, the issue forced a vote and the count was one vote shy from forcing a change in leadership. In 2014 and 2015 another hacktivist, Phineas Phisher, hacked in the service of data leaks by striking against two firms, Gamma Group and Hacking Team—firms suspected of selling surveillance software to totalitarian regimes. Like previous cases discussed here, his liberation of Hacking Team's emails served as an evidential anchor by confirming suspected wrongdoing. This was put well to me by Lorenzo Franceschi-Bicchierai who covered the hack for VICE Motherboard: "Before Phineas Fisher broke into the servers of Hacking Team, we already suspected, based on extensive and detailed research, that they were selling [spyware] to oppressive regimes. But his hack gave us the ultimate proof."[2] A year after the hack, Hacking Team lost its license to export spyware outside of the EU.

Up until 2014, public interest hacks were solely the domain of hacktivists. But in the summer of 2014, a distinct and more mysterious species of hacker would deploy this tactic. Unlike hacktivists who transparently express their objectives, these actors advertised their hacks, but never disclosed their true intent.

The first hack to unfurl in this new guise struck like a tempest in 2014, when a mysterious hacker group, Guardians of Peace (GOP), ransacked and pillaged Sony's servers, dropping company emails into the public. It was an attack characterized by security and government officials as "unprecedented"—largely, I would suggest, for its PIH characteristics. Eventually the GOP specified that their actions were taken in vengeance for a Hollywood film—*The Interview*—that poked fun at the North Korean dictator. The journalistic analysis, which was gargantuan, largely concentrated on the intrusion, extortion, motivations, and forensics of the hack rather than the content of the emails. Still, some journalists excavated the material for salacious gossip about celebrities written by executives, while others used it for social commentary: uncovering disparities in earnings by gender and race. What was already known was made explicit, with exact financial figures suddenly made available.

While the US government blamed the North Korean government, the hack baffled many security experts; some of whom insisted the claim rested on shaky, inconclusive evidence (Zetter 2014). Determining whether or not the North Korean government masterminded the hack or only later piggybacked on its coattails may prove unimportant; this hack offered another public statement that conveyed in effect that a government or other entity *could* use this method for a motley array of purposes, such as retribution, a raw display of aggression and power, or other geopolitical machinations.

It appears that at least one powerful nation has since heeded the lesson. Nearly two years later, a similar hack—similar insofar as the ulterior motive was concealed—was leveled against the Democratic National Committee (DNC), leading to the disclosure of multiple email caches. The hacker-in-chief laying claim to intrusion went by Guccifer 2.0. In contrast to the Sony hack, different sectors of the security community were nearly unanimous in their assessment: everything about the hack—from forensic to other geopolitical evidence—pointed to Russian intelligence.

This hack and leak leapfrogged past the GOP Sony hack to become the single most controversial PIH to date. The fallout from the hack was volcanic, with raging disputes spewing to this day about its source, impact, and meaning: scores of liberals were dismayed that the emails might have thwarted Hilary Clinton election bid; Bernie Sander's supporters were livid that the correspondence demonstrated the DNC failed to play fair; and some pundits and journalists harrumphed that the emails contained no meaningful material whatsoever (see Sauter, Colvin, Fish and Follis, and Gorham in this issue for contrasting takes). Some information liberation advocates were upset that WikiLeaks chose to publish the emails at all, while others supported the embattled organization—asserting that truth is not distorted by its messengers. Elsewhere, various pundits: wished the material had been published only after the election; forecasted the start of new cold war with then President Obama shortly thereafter booting thirty-five Russian diplomats from the US; maintained the Russian hysteria was overly-hysterical; and used the emails as raw ingredients to cook up the dangerously weird conspiracy theory, Pizzagate.

The DNC hack/leak, thoroughly defined along numerous fault lines, unfurled over time with divergent consequences. The DNC emails were used by some journalists to break stories. But the material could also be used to unleash a thicket of confusion or, what might be better called (with a nod to the fog of war) the fog of hacking—a hack and leak designed to distract, confuse, and seed doubt in the public.

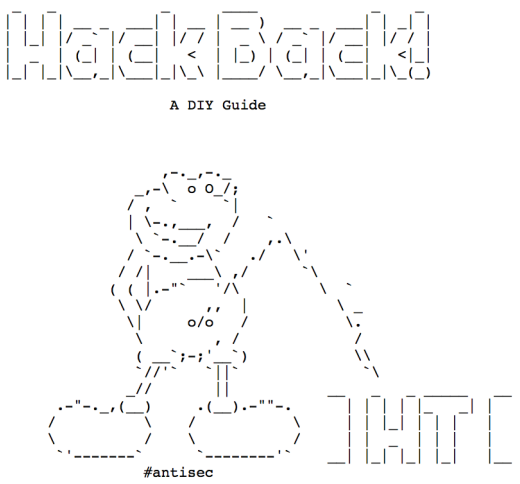
CONCLUSION

The history of the PIH may be remarkably recent but it seems here to stay. Indeed, 2017 has already seen a number of high-profile instances, such as the hack of Cellebrite, an Israeli mobile form, with the hacker first channeling some documents directly to a journalist and subsequently publicly dumping the firm's circumvention tools. Another even more notable example is the gargantuan hack against Retina and FlexiSpy—software companies marketing "stalkerware" to other firms and individuals for monitoring employees or children. Entering and then swiping source code, HR documents, and other files, the hackers leaked this information, which became the basis for a series of investigative pieces detailing how this spying software is used by "lawyers, teachers, construction workers, parents, jealous lovers" (Franceshi-Bicchierai and Cox 2017). Clearly following the path blazed by hacktivist predecessors, these hackers, going by name the Decepticons, also published a "How-to guide for aspiring hackers" with a respectful shout out to Phineas Phisher, noting: "we'd be remiss if we didn't include Phineas Phisher's articles, which are fantastic introductions. They cover things like how to stay safe and many of the basics, including many techniques we used to compromise FlexiSpy/Vervata/etc. So read them and soak them up" (Decepticons, 2017).

Some might be wondering whether the Shadow Brokers' April 2017 dump of NSA hacking tools qualifies as a PIH under the rubric proposed here. Given available information, it's hard to say. Journalists certainly mined the leaked data and tools to unveil

1 A more comprehensive history of the PIH would also need to examine other regions, such as Asia and the Middle East and especially Turkey home to a prolific hacktivist group, RedHack.

2 Personal Communication with the author.



Artwork from Phineas Phisher's "HackBack: A DIY Guide."

new details about the use of exploits and malware by US intelligence, but evidence as to whether the data was acquired from a hack or by some other means remains circumstantial. According to Edward Snowden, this group—likely composed of nation state-backed hackers—infiltrated a staging server (itself a hacked server where the NSA would host and launch their tools) where they discovered the tools left for the taking. This hack would not be "unprecedented." But what is unprecedented is the publicness, the style of "publication," as Snowden put it, of the material (2016).

That there is a connecting thread between Anonymous, Phineas Phisher, and the Decepticons is obvious, confirmed by the actors themselves—each subsequent hacktivist paying homage to their predecessor. In contrast, it is impossible to say definitively whether groups like Guardians of Peace, Guccifer 2.0, or Shadow Brokers were overtly or directly influenced by Anonymous. What is evident—and the recent hack and leak of Macron staff email provides another nugget of proof—is hackers will continue to rely on but also experiment with this method. And experimentation invariably leads to mutations. The PIH will continue to be used as it has been in the last few years: as an instrument for left-leaning hacktivism, statecraft, revenge and extortion, and geopolitical machinations; but as journalists develop new norms for reporting on leaks and as hackers become more sophisticated at launching and staging attacks—for instance, by successfully implanting false information in the leaks—the form will continue to surprise us with its myriad political effects and consequences. ■

E. GABRIELLA COLEMAN *holds the Wolfe Chair in Scientific and Technological Literacy at McGill University. She is the author of* Coding Freedom: The Ethics and Aesthetics of Hacking *(Princeton University Press, 2012) and* Hacker, Hoaxer, Whistleblower, Spy: The Many Faces of Anonymous *(Verso, 2014).*

BIBLIOGRAPHY

Anderson, Nate. 2010. "The 'legal blackmail; business: inside a P2P settlement factory." *Ars Technica*, September 29. Available at link.

Berman, Dennis K. 2003. "Government Posts Enron's E-Mail: Amid Power-Market Minutiae, Many Personal Notes Remain." *Wall Street Journal*, October 6. Available at link.

Decepticons. 2017. "Flexidie: brought to you by LeopardBoy and the Decepticons." *Pastebin.* Available at link.

Franceshi-Bicchierai, Lorenzo and Joseph Cox. 2017. "Inside the 'Stalkerware' Surveillance Market, Where Ordinary People Tap Each Other's Phones." *MotherBoard*, April 18 Available at link.

Grieve, Tim. 2003. "The decline and fall of the Enron empire." *Salon*, October 14. Available at link.

Harris, Bev, 2003. *Black Box Voting Book.* http://blackboxvoting.org/black-box-voting-book/

Menn, Joseph. 2011. "Cyberactivists warned of arrest." *Financial Times*, February 4. Available at link.

Phisher, Phineas. 2016. "Hack Back! A DIY Guide." *Pastebin.* Available at link.

Potok, Mark. "Neo-Nazi Threatmaker Accused of Working for FBI." *Southern Poverty Law Center Hatewatch*, January 11. Available at link.

Schwartz, John. 2003. "Computer Voting Is Open to Easy Fraud, Experts Say." *New York Times*, July 24. Available at link.

Schor, Elana. 2008. "US election: Tennessee politician's son indicted for hacking into Palin's email." *The Guardian*, October 8. Available at link.

Snowden, Edward (Snowden). "The hack of an NSA malware staging server is not unprecedented, but the publication of the take is. Here's what you need to know: (1/x)". 16 Aug 2016, 11:40 UTC. Tweet. Available at link.

Zetter, Kim. 2014. "Experts Are Still Divided on Whether North Korea Is Behind Sony Attack." *WIRED*, December 23. Available at link.