International Journal of Police Science
ISSN 2771-3377 (online); ISSN 2771-3385 (print)
https://IJPS-Journal.org
Volume 1, Issue No. 1 (2022)

International Police Science Association (IPSA)

Review Article

How Online Information Sources Affect Law Enforcement Perceptions

David C. Benson

https://doi.org/10.56331/487529/ipsa1

Published: 26 June 2022
Corrected: 10 April 2023*

Citation: Benson, David C. "How Online Information Sources Affect Law Enforcement Perceptions." *International Journal of Police Science* 1, no. 1 (2022). https://doi.org/10.56331/487529/ipsa1.

Abstract: The internet is a part of daily life, but the law enforcement community is still struggling to adapt policing research and practice to the internet due to a limited understanding of the various online information sources affecting law enforcement and security. This article addresses information as a function of policing or crime and focuses on how online sources of information may be used to prepare for, commit, prevent, investigate, and prosecute crimes.

Keywords: internet; surveillance; information security; data analytics

Introduction

In most of the world, the internet is a part of daily life, but legal systems and law enforcement practices are still adapting to the internet. One stumbling block to adapting policing research and practice to the internet is a limited understanding of the various online information sources affecting law enforcement and security. The internet is an information technology. Many laws governing information and crime were developed under assumptions about information that may no longer hold. Incorrect assumptions about information can frustrate police and criminologists as they work to apply existing, otherwise valid, theories and practices in the new information environment. Legal systems and law enforcement need to adapt their extant laws and policing practices to the internet's novel information topography.

This article explains new sources of online information law enforcement and criminology must account for by explaining the three vectors of information, unique to the internet, that affect public security. The taxonomy this article explains focuses on real-world, observable behaviors. In other words, this taxonomy creates categories based upon observable behaviors, not information technology. Although other technological taxonomies are referred to where appropriate, information taxonomies reliant on technology do not map neatly onto real-world behaviors. Some information vectors can use different underlying technologies to create the same real-world outcome, while others use the same technology for different real-world outcomes. For example, both phishing and vishing trick users into accidentally surrendering their logins, but phishing uses email while vishing happens over the telephone.[1] Similarly, criminals and companies both harvest online data using web scraping, but criminals use their web-scraped data to commit crimes, while companies use their information to build market-share.[2]

This article organizes information collection into three groups based on how information collectors acquire the information: 1) public information aggregation; 2) deception, coercion, and inducements; and 3) information security violations (hacking). Each category poses challenges and opportunities to law enforcement. No category maps neatly onto criminal behavior as defined in most legal systems. Laws change from place to place and from time to time. Any taxonomy that perfectly reflected a specific legal structure might lose its value when laws change or might be worthless outside a specific jurisdiction. This taxonomy allows professional police and criminologists to evaluate information-relevant practices without resorting to in-depth knowledge of the base technology.

Different information sources have different security implications. Criminals using information for crimes is an inherent problem, regardless of how the criminals acquired the information. How police can interdict or exploit criminal information acquisition and use will depend on the information source. Grouping information in a way that allows appropriate analytical comparisons enables developing appropriate understandings of the new information technology. Criminals collecting information from public directories is a materially different problem from criminals who access confidential information in a private database.

Police and criminologists must account for the new information logics; others cannot do it for them. Even if information professionals comprehend online information, information professionals

do not have the depth of knowledge or experience to apply the internet's information logics to police science. At best, adjacent researchers may provide useful analogs that criminology could apply. Scholars across disciplines must communicate with each other, but experts must still integrate outside ideas into their calculations. Just as it is unfair to expect criminologists to develop a keen understanding of information technology, it is also unfair to expect that information experts will understand criminology.

This article addresses information as a function of policing or crime and leaves aside other roles information could place in normal policing. Police may also order food online during late night work, but such economic use parallels the ways non-police use the internet. While interesting questions unto themselves, the purpose of police science is to understand crime and law enforcement. Therefore, this article focuses on different information sources and how they may be used to prepare for, commit, prevent, investigate, and prosecute crimes.

Discussion

Cyberspace as a Policing Space

Scholars and practitioners of police science should resist the understandable tendency to think of cyberspace as primarily "the place where cybercrime occurs." Many policymakers tend to view policy questions primarily as challenges to overcome, or threats to dispense with. The pessimistic tendency is especially strong in professional communities dedicated to security, like criminology, because the profession exists to provide security. There is a natural tendency for national security professionals to view cyberspace as a place for war, for cybersecurity professionals to view cyberspace as cybersecurity vulnerabilities, and for criminologists to see cyberspace as the source of cybercrime.[3] Cybercrime happens primarily in cyberspace, of course, but cyberspace is not primarily a place for crime.

Cyberspace as a Commons

Cyberspace is a space for human activity of all sorts, of which crime is a small part. Information passes through, inhabits, and defines cyberspace. Any activity using information will eventually make its way online when people supplement traditional information technologies with the internet for communication. Consequently, the internet hosts crime, but also social interaction, commerce, and competition.

Policing provides security, but police can use cyberspace and information found therein beyond merely providing security. Police are expected to help protect information from criminals, but even criminals may use otherwise legal information in crimes. The information that may matter in crime prevention may not include information protected under the law. Real-world crimes may use online information and online crimes may use real-world information or happen primarily in the real-world. Police can use online information in law enforcement, just as all other people use the internet in their lives.

It may be useful to think of the internet as a constable in a nineteenth century small town might have thought of the town square. The townsfolk would expect the constable to prevent theft in the town square—especially on market days—but if the constable approached the square primarily as a policing challenge, the constable would do more harm than good. Observing people,

especially potential ne'er-do-wells, might tip the constable to upcoming crimes like bank robberies. Speaking with townsfolk about their concerns provides information on the state of the town, but also builds relationships that may be valuable in future investigations. People visiting the square might be there to shop in a store or at a fair, and the constable might also purchase goods that way. Smart constables would take crime in a public space seriously but would also use it just as anyone else would use it.

For a small-town constable, the town would also be a partner in providing its security. The law-abiding townspeople have just as much interest in protecting their security as the constable the town employs. Townsfolk need not be as martially-minded as Northfield, Minnesota, who famously fought off the James-Younger gang's attempted bank robbery to take measures to defend themselves.[4] Locking doors, reporting malefactors, and general personal security practices all help the constable's work. Members of the public participate in their security by cooperating with police. Many of the institutions at the forefront of online security are non-governmental. Even international terrorists online find both police and civilians cooperating to thwart terroristic objectives.[5]

Technology Changes Caused Information Changes

The internet's underlying technology operates under a fundamentally different logic from legacy information technologies. Information still flows over the internet, much as it does over other information technologies, but the way the internet passes information changes the ways people can get information. Some of the logics that were true under legacy information technology remain true with the internet, but some do not. Strictly using legacy information logics to think about online information will be fruitless; creating novel logics where legacy logics apply wastes time.

The internet's main innovation in information technology is the use of self-routing packets across an interconnected network. The Transfer Control Protocol/Internet Protocol (TCP/IP) breaks information into packets that independently make their way through switches to their destination. Self-routing packets disconnect information passage from the underlying geography. No two packets need to take the same pathway from source to destination, making the network itself reliable. When recipients acknowledge packet receipt, information transmitted online becomes among the most reliable modes of communication.[6]

Computers enable the internet, and the internet makes computing more important for information. Information can be widely available when online because reliably replicating information at multiple locations is comparatively cheap.[7] Digital information allows both simple and complex digital information processing techniques like machine learning.[8] Computers are not inherently secure against information theft, and stolen information of previously inconceivable quantity and quality is increasingly a facet and source of online information.[9]

International Internet

The internet can bring all the world, for good and ill, into each jurisdiction. Before the internet, transnational crime was usually limited to a few kinds of crimes or locations. Transnational criminals might plausibly bedevil major ports or afflict currency trading, for example, but would rarely if ever be a challenge to the average citizen. It was possible to have mail fraud across

international boundaries, but international mail was costly, making defrauding laborers and housekeepers unprofitable. Broadcasters could shill snake oil from transnational radio stations on borders, but those stations' reaches were limited and could be traced. Outside of national police, customs agents, and intelligence organizations, police rarely encountered transnational crime. Local police could rely upon the national government.

Online transnational crime now touches every jurisdiction and can impose tremendous costs. In 2020, cybercrime cost the global economy at least $1 trillion, or approximately 1 percent of global GDP.[10] According to early estimates, cybercrime in 2021 could total $3 trillion and could reach $10 trillion as soon as 2025, unless trends change.[11] While high-profile cybercrime tends to afflict large, multinational institutions, smaller organizations and individuals make up the most frequent victims. In 2021, cybercrime attacked victims as far-flung as schools in Janesville, Wisconsin; hospitals in Waikato, New Zealand; and banks in Lagos, Nigeria.[12] Even crimes with long histories in local jurisdictions, like stalking, can assume new nefarious characteristics when merged with online information.[13]

Police may also find that international criminals are not behind cybercrime, but other governments' agents. Many of the most capable and pernicious advanced persistent threats (APTs) operating are government-sponsored. When a group attempted to steal nearly $1 billion (successfully stealing $80 million) from the National Bank of Bangladesh using Society for Worldwide Interbank Financial Telecommunications (SWIFT) transfers, analysis indicated a North Korean, government-supported APT commonly called the Lazarus Group was responsible.[14] Investigators believe that a People's Republic of China (PRC) government-sponsored APT that Microsoft calls Hafnium was behind the attack on Microsoft Exchange On-Premises servers that impacted tens of thousands of mostly small businesses in 2020.[15] The Russian APT that created the costly and sophisticated custom tools used in the SolarWinds hack revealed in late 2020 were probably government supported.[16]

The potential for a foreign government to be responsible for cybercrime increases the complexity of protecting against cybercrime. When interrupting a home break-in, police are unlikely to discover the intruder is a uniformed member of another government's military, but an equivalent outcome is possible, and sometimes even likely, online. Foreign government agents have resources, including international protection from prosecution, that most criminals do not.[17] Law enforcement may need to account for the effects of international diplomacy on both their investigations and their ability to protect potential victims in their jurisdiction. Law enforcement has had to deal with international involvement before, such as the German police investigation that revealed a Russian government plot.[18] Cybercrime increases the likelihood that police might investigate a crime implicating a foreign government, with all the complexity such an implication may entail.

Information Security Breaches (Hacking)

An information security breach, or "hacking," is probably what most people think of when they think of the challenge the internet poses to law enforcement. I use hacking in this article to mean breaching one or more information security protections. Hacking has other meanings inside and outside information, and some writers even use it to describe the information collection vectors described above. In most common parlance and this article, information security hacking implies a high degree of technological sophistication.

Hacking Exploits Security Weaknesses

Hacking exploits technology weaknesses, including flaws in coding, configuration, and implementation. Vulnerability exploitation includes vulnerabilities introduced by users themselves and technological flaws in code known as zero-days, worms, trojans, and viruses. Vulnerability exploitation sometimes requires special knowledge and may differ substantially from traditional information collection. Vulnerability exploitation also requires no witting interaction with approved users. Hackers can exploit vulnerabilities to collect information with reduced risk of alerting the target. Thus, cybersecurity vulnerability exploitation appeals to actors attempting to collect information clandestinely.

"Zero-day vulnerabilities" exploit weaknesses in computer code. Zero-days are probably closest to what is popularly portrayed as hacking in fiction. Attackers can exploit zero-day vulnerabilities despite the victim's best efforts. Flawless cybersecurity hygiene and maximal defensive posture cannot defend against weaknesses in the programs you use. Groups or individuals wanting to use zero-days can employ researchers to look for zero-days, or purchase zero-days on the open market.[19]

"Hacking as a service" allows actors to purchase capabilities they do not possess themselves.[20] Researchers only selling zero-days for legitimate purposes are called "white hats." "Black hats" will sell exploits to the highest bidder. Hackers can also learn many necessary skills for information security vulnerability exploitation through zero-day markets.[21] Increasingly, it is possible to simply treat hacking like it were any other service and purchase it, allowing states to simply buy with money a capability they have not developed themselves.[22]

Users create vulnerabilities by using poor information security hygiene, like not having a password and improper configuration of security protocols. Even using weak or common passwords make users vulnerable to exploitation, because malicious actors can guess the password easily. As late as 2017, the two most common passwords were "123456" and "password," which does not provide much security.[23] Recommendations against using birth dates, names, names of family members, or other closely associated words are intended to prevent people from creating easily guessable passwords. Reusing passwords is nearly epidemic.[24] Even when available, people will opt not to use security-enhancing features like multi-factor authentication.[25] Computing power makes commonly reused passwords incredibly vulnerable to discovery, even if hackers do not publish the passwords.

Sophisticated hackers can develop malware, which exploits either individuals or software to introduce vulnerabilities to a system. Some malware, such as "China Chopper," exist as tools that hackers can use like thieves might use a crowbar.[26] Other malware, such as "Sunburst" or "Stuxnet," are custom creations made for individual targets, though potentially useful on other targets.[27] In most cases, an unwitting stooge installs malware unaware of its real purpose, but sometimes hackers can introduce malware directly, such as through software supply chains.[28] Ransomware is a type of malware that has become recently pernicious, costly, and effective for criminal enterprises. Ransomware introduces a program to the victim's computer that password-encrypts the victim's data. If victims pay a ransom, the hackers will provide the decryption password for the encrypted data.[29] In some instances, the hackers may leak stolen information

if the victim does not pay the ransom.[30] Over the past several years, ransomware has become a rampant problem and has even been implicated in two deaths.[31]

Criminals Use Hacking

Hacking is appealing to criminals because hacking often accesses information without the information owner's knowledge. Exploiting a zero-day vulnerability allows hackers access to privileged information, no matter what information defenders may do. Unless the administrators recognize that a breach has occurred through other means, the breach could well pass completely undetected. The attackers who infiltrated the Democratic National Committee did so by exploiting several zero-day attacks and were not found out for months.[32] Malware victims must usually install malicious software but usually do so without recognizing that the software they are installing is malware. In either case, hackers good and malicious operate under a temporary cloak of secrecy, that must be deliberately lifted allowing them time to slip away from detection.

Hacking by itself can be very profitable. Exploiting information security vulnerabilities has reliable mechanisms that pay hackers. Information is inherently valuable to individuals and companies, and unsurprisingly they will pay hackers to return it. Many companies even take out insurance which will pay hackers for them, specifically to guarantee they will be able to pay criminals.[33] The owners of the Colonial Pipeline paid $4.4 million ransom to have their data restored.[34] In some cases, hackers can even transfer money directly to their bank accounts. Hackers were able to steal more than $80 million directly from a bank in Bangladesh.[35]

Hacking can also acquire information that allows other profitable crimes. Stealing a credit card number or personal information can be profitable because criminals will purchase that information, but such information is useful because it can be used in criminal enterprises to get money. Hackers stole payment information from Home Depot and Target, and then were able to use that stolen information to steal money from secondary victims.[36] The credit data hackers stole from Equifax were primarily useful because the data could enable either the hackers or their clients to steal identities.[37] Criminals may have motives other than profit but profit alone is sufficient to ensure that cyberattacks for criminal enterprises will continue.

Police and the Public Use Hacking

Police use many of the same hacking tools for policing as criminals. Just as a wiretap is technologically identical whether used for fraud or police surveillance, exploiting malware and zero-days are identical whether used for licit or illicit purposes. What determines whether any hacking is legal or not is whether the information collectors followed legal processes when collecting information. When the FBI wanted to access encrypted iPhones used by the San Bernardino shooters, it turned to a company specializing in "white hat" hacking.[38] The FBI collaborated with criminals and law enforcement around the world to seed criminal networks with a corrupted messaging service to observe criminal communications.[39] In both cases, law enforcement behavior not only overlapped with crimes but was only distinguishable by the legal sanction that proper police procedure provided.

The overlap between criminal and policing hacking is likely to be a consistent point of contention complicating cooperation among police departments. If police, intelligence, and white hat hackers develop tools for legitimate reasons, those tools can be misused for criminal purposes. Security

researchers originally developed *Cobalt Strike* to test and improve security, but it has since become a tool in criminals' tool chests, too.[40] Some governments may develop and use tools to target groups that not every government agrees is criminal, creating international friction. The Chinese central government considers some groups like Uighur and Tibetan activists criminal organizations and uses hacking tools to surveil those groups.[41] Whether tools or targets create the contention, police and police science must actively grapple with this novel form of information collection because it is so unlike anything before.

Law enforcement can partner with other groups working to ensure security. Many companies and organizations exist to defend cybersecurity, and frequently better understand cybersecurity than many police. Security researchers called "white hats" look for information security vulnerabilities to fix those weaknesses, exactly like criminals—"black hats"—who use vulnerabilities to commit crimes.[42] Large technology companies like Microsoft, Apple, and Google all maintain cybersecurity researchers. Smaller companies exist exclusively to provide cybersecurity services. Over-rigorous prosecution of legal prohibitions can harm information security by reducing domestic security capabilities.[43]

Deception, Coercion, Inducement

The simplest, though not necessarily the easiest, way to access confidential information is to have someone with access to confidential information give it to you. Many of the mechanisms people have used for ages off-line work just as well online. Non-public, confidential information is an attractive target for both criminals and law enforcement. Criminals can use confidential information in the commission of crimes. Possessing confidential information without authorization might be a crime itself. Law enforcement protects the public's confidential information and acquires confidential information about crimes to enforce the law. Both criminals and law enforcement naturally adapt extant techniques to acquire confidential information.

Getting Information by Lying, Threatening, or Bribing

If someone has something you want, it can often be easier to get the person to give you the thing you want than try to take that thing from them. Sometimes the only way to get information is to have someone who has information give the information to you. One way to get someone to give you the information you want is to trick them or lie to them, but you must be able to deceive the information holder. Threatening people or bribing them can work, although threats and bribes rely on your ability to carry out the threat or the bribe. In either case, the individual without the information passes information in the form of a lie, threat, or promise of a bribe, in exchange for the desired information.

Online deception, coercion, and inducement simply port off-line actions to great effect. Impersonating a person who has access to confidential information is a way to "social engineer" a hack.[44] Inculcating oneself into a social media group by impersonating someone else can also give illicit access to information.[45] Insiders who wittingly provide information to illicit sources, recruited either through threats or bribes, are a major threat in information security.[46] Regardless of the mechanism one uses online, the result is that an authorized information holder provides the confidential information access to an unauthorized individual.

Deception, coercion, and inducement online use the same logics online as offline but differ in important particulars. Deception online is easier, mostly because careful deceivers can greatly reduce the amount of information shared with their targets. It is easier to pretend to be someone you are not if the only form of communication is written text. Tools and capabilities exist to determine identities, but even if people have the expertise to use those tools, part of deception is instilling in the target the confidence that they need not investigate further. Coercion and inducement strictly online can be more challenging because many of the threats or promises one would make require personal contact. Coercion and inducement may eventually result in real-world interaction. However, sometimes threats are unique to the online environment, like blackmail using illicitly acquired photographs.[47]

Criminals Use Deception, Coercion, and Inducement

Criminals use online deception to commit fraud. Posing as a potential love interest, or otherwise exploiting prurient appetites, has a long history within espionage, known as the "honeypot" or "honey trap" and is called "catfishing" online.[48] Online dating has made romance scams even easier, by allowing individuals access to potential victims globally, pre-sorted according to preferences and location.[49] Real-world romance scammers had to be suave, or good looking, and at least present. Online romance scammers can attack all around the world, using someone else's image.

Criminals can also use deception to acquire access to other information. "Phishing" and the more targeted "spear-phishing" are technological forms of dissembling whereby one acquires electronic credentials by misrepresenting oneself to a target. In most cases, spear phishers present targets with an opportunity to provide their credentials—their login, password, etc.—to a seemingly legitimate portal.[50] In reality, once the user provides their credentials, the portal will pass them to malicious actors intent on gaining access to restricted information. Phishing emails may be ham-fisted, and most internet users will receive several phishing emails as spam. Spear-phishing is sophisticated, often including personalized information and appearing to come from a known associate.[51]

Coercion and inducement online produce "insider threats."[52] Insider threats are among the most devastating information security vulnerabilities because insiders are supposed to have access to that information. A typical insider threat would be the thwarted attack on Telsa. Egor Igorevich Kriuchkov offered a Tesla employee $1 million to emplace ransomware, but the employee immediately informed on the hacker.[53] Had the attack succeeded, it might have been among the most costly attacks ever. Among the most famous, and damaging, national security leaks were those released to Wikileaks by Snowden and Manning.[54]

Police Can Use Deception, Coercion, and Inducement

Because there is contact between victims and criminals, deception, coercion, and inducement offer obvious starting points for investigation. Once someone is alerted to their deception, they at least know they have been deceived, and the criminal may leave forensic profiles. Investigators can often identify insider threats, and use those as a starting point for their investigation. When "hackers" tricked a Twitter employee into disclosing a password and took control of high-profile Twitter accounts, police used the initial contact to find and arrest the three hackers responsible.[55]

Furthermore, sometimes deception, coercion, and inducement themselves are inherently illegal, allowing enforcement to focus on information collection rather than subsequent crimes. In most cases, one need only know an assailant is threatening a victim to know that a crime is being committed. It is still criminal to threaten to blow someone's brains out with a pistol you are brandishing unless they give you their password, even if you only want the password to check sports scores. Inducement, or bribery, is illegal in many jurisdictions, though enforcement of anti-corruption statutes varies across jurisdictions.[56]

Policing can also use coercion, deception, and inducement as a part of legal processes, but in a different form. Warrants compel compliance by threatening imprisonment; police induce informants with payment; undercover operations are deception. Lessons learned during investigations help to better understand parallel criminal processes.

While police and criminals operate under different constraints, when information acquisition is difficult for police it may be equally difficult for criminals. If the information police receive from warrants is unusably complex or vague, it is plausible that criminals will find the same information equally unusable when obtained illegally. Requiring companies to maintain records suitable for police use also ensures that, if stolen, the information will be equally useful for criminals.

Mass Public Data Collection

While data collection has long been socio-politically important, the internet allows researchers, analysts, and investigators to collect and analyze data at a previously inconceivable scale. Writers call mass data collection and analysis big data or data analytics in different contexts.[57] In this article, I treat big data and data analytics as roughly coterminous, although they take on different meanings. Strictly speaking, big data focuses on the available data quantity, whereas data analytics' focus is the techniques required to distill useful information from those data. From the end-user perspective, such as law enforcement, the mechanisms at play are of less importance than the resulting information.

Big Data Collects and Analyzes Publicly Available Data

Big data and data analytics are theoretically possible without the internet, and non-automated data collection can yield useful information. Employers checking potential employees' social media is a common form of manual public data collection.[58] Slightly more technical solutions include companies and services like IFTTT, Zapier, or Microsoft Power BI. As data science develops, increasingly capable companies and consultants are developing purpose-built data collection tools. Almost everyone online can access big data without necessarily knowing much about the information collection and analysis project.[59]

The internet makes big data collection and data analytics practical, and the insights from big data can be profound. TCP/IP makes automated information collection—data scraping—easy, compared to data collection with legacy information technologies. The only resources necessary to collect online information are a computer, an internet connection, and any freely available tools. For example, with nothing more than a cheap laptop, a researcher can scrape all of the tweets on any subject taking only as long as the download.[60] It is tautological, but not trivial, to say that members of the public can collect all information available publicly online. Both individuals

and observers seem to forget how much information is already publicly available and fail to adjust their behavior accordingly.

Online data analytics can both distill information about individuals and extrapolate information about the broader community. For example, searching publicly available photographs on social media, a publicly available program can identify someone's physical location.[61] Other researchers demonstrated that they could identify homosexuals using their public social media profile, even when the individuals did not self-identify.[62] At the other end of the spectrum, data from search engines can track influenza outbreaks.[63] Aggregate information can potentially predict seemingly inscrutable events.[64] Using data analytics to aggregate public information into useful aggregate data parallels the ability of markets to aggregate private information in the form of price.[65] During the COVID-19 outbreaks, analysts could use data from internet-enabled smart thermometers to predict where COVID-19 risk was increasing, even before traditional tracking by the US Centers for Disease Control had identified increased risk.[66]

Although data analytics can be powerful, the information analysts will discover using big data analysis is not always obvious ex ante. The relatively inchoate state of data analytics accounts for some of the unpredictability of big data results. As the field develops and diffuses, experts will be able to better predict what insights large, publicly available data can produce. Data analysts will likely always be surprised by what they cannot discern because big data collection is largely "catch-as-catch-can." If analysts cannot acquire the appropriate data or lack a way to analyze the data, then they will be unable to discern what they had hoped for. On the other hand, we are also likely to be constantly surprised at what big data can discover. With the online environment awash in data, enterprising analysts need only discover new measures and suitable analyses to discover previously hidden information. Artificial intelligence and machine learning (AI/ML) further increase the likelihood of novel information discovery.[67]

Consequently, police science and society at large must constantly be on guard for possible ways to exploit big data because we simply do not know which avenues will prove threatening.
Many companies use publicly collected data and data analytics to better tailor their business strategies, and law enforcement may be able to avail themselves of similar capabilities. Market research has probably been a part of business since the dawn of markets, but the internet's development enables increasingly sophisticated and data-rich analysis.[68] Government data sources can be among the most important sources for businesses in their data analytics.[69] Companies use their analytics to fit their offerings into the extant market, including accounting for consumer demand and competitors' strategies.

Entities can collect public information without the target's knowledge or explicit consent. Many online services offer an application programming interface (API) that allows programmatic access to the service's information. Collecting information using APIs imposes some constraints on info collection, but rarely notifies information users.[70] Cambridge Analytica's information collection scheme using Facebook's API neither violated Facebook's official terms of service nor notified information owners that Cambridge Analytica had collected their information.[71]

Website scraping can collect any information posted on the world wide web without an API. Services can collect information from websites, including news publications or personal websites, for both licit and illicit ends.[72] Laws like the EU's General Data Protection Regulation (GDPR) limit what information services can store, hypothetically reducing information exposure.[73] Limiting

information storage may reduce the information available publicly, but does not reduce the ability to collect information.

Collecting public data online usually either collects from a list of sources or a "walk" of links or references. Scraping data following a list involves visiting a delimited set of data sources and collecting all the information found there, like a list of monitored websites or social media feeds. Walks start from a list, and then follow an algorithm at each site following links to other information sources, like starting with a Twitter feed and then scraping all feeds mentioning the initial feed, and each feed mentioning subsequent feeds, etc. List-based data scraping efficiently keeps track of known data sources but rarely encounters new data sources, creating blind spots. Walk-based data can find new data sources, but following walks can be slow, and not all the information sources the walk finds will be useful.

Criminals Can Use Big Data

Knowing how criminals use big data can be challenging and may not necessarily appear in the public record. In most jurisdictions, using publicly available data is not criminal. Even if big data played a crucial role in a criminal enterprise, it may be unimportant or even a distraction to criminal prosecution, and may not appear as part of indictments. Bank robbers might use big data to identify good targets, but if an observant beat cop captured them, the police might be unaware of the role big data played in the operation. Furthermore, prosecutors could decide—probably correctly—that introducing statistical analysis to a jury would impair prosecution, even if the police knew of the criminals' data analytics.

Because of information limitations, speculation and extrapolation play a larger role in understanding how criminals use big data. The reliance on speculation is not necessarily problematic, however, as it encourages criminologists to be imaginative. A forward-thinking investigator may be able to identify potential criminal uses of big data and head them off before it becomes a problem. Identifying potential uses before we see actual uses can also establish measures that would identify when criminals begin to use big data. Ultimately, either police science or criminals will come up with a use case; the only question is who gets there first, and what the effects are on society.

Police science can rely upon examples of big data use as starting points. Using legally acquired information can allow criminals to tailor their crimes toward specific victims. For example, the criminals who stole John Podesta's personal emails relied on public information to identify potential targets.[74] Infamous "hacker" Guccifer never hacked his victims. Instead, he guessed passwords using their public information.[75] Criminals also identify targets or times for real-world crimes, such as identifying when potential victims are not home, using social media.[76]

Examples of non-criminal big data use can also indicate how criminals might use big data by identifying licit big data uses that could become illicit in different contexts. The identification of potential victims could use similar data and analytic techniques as the identification of potential customers. We may have already seen such uses of big data by criminals during the COVID-19 pandemic. During the pandemic, law enforcement observed a spike in fraud related to COVID-19.[77] Although one does not need big data to identify COVID-19 as a major issue, it is not difficult to imagine that criminals are using publicly available data to identify good targets for their crimes. COVID-19 caused different behavioral changes in different groups, each with their identifiers, and

with their vulnerabilities for crime.[78] Analysis and investigation may yet reveal that criminals were using many of the same data sources and analytic techniques to identify their victims as doctors and public health officials were using to identify health needs.

Police and the Public Can Use Big Data

In most jurisdictions, collecting and analyzing public information is not illegal, but could be used both in criminal and policing functions. Acquiring inherently public information is a natural part of daily life and would be difficult to regulate. Even searching for misconfigured server passwords that would be easy to exploit can be legal and useful to find security vulnerabilities. Acquiring public information online will consequently look similar, whether the outcome is criminal or legal. Consequently, criminal and police use of public information will also parallel one another.

Police can use public information to more easily thwart individual crimes and criminal trends. The ability of police to collect information about specific criminals will vary by jurisdiction. Even in jurisdictions where legal rights constrain police collection, civil informants are often willing to use public information to provide information to the police. Police can also follow the same aggregate information as criminals and use that information to preemptively identify potential vectors for crimes. If police can identify neighborhoods where many homes will be vacant over the holidays, criminals probably can too, and police can increase patrols in that area to preempt potential crime.

Data analytics are powerful tools for the public to ensure its security. Many of the same trends the police identify would be available to the public for analysis. Private citizens, groups, and organizations can also use their privately collected data and share insights with the police. In the same way, stores cooperate with police to tamp down theft and credit card fraud at the local level, they can cooperate to identify risks to public safety in their data. Many such endeavors exist already in the developed world, but as the internet increasingly infiltrates daily life those opportunities will expand.

Police using data analytics may rouse popular objections and may pose constitutional challenges in many countries. Although some case law and legal protections exist, no consensus has emerged on which kinds of public information police can use. Following *Wikileaks'* publication of documents Edward Snowden exfiltrated, a firestorm erupted surrounding government use of cellphone metadata, one source of big data.[79] Although Snowden's leaks dealt with National Security Agency programs, (an intelligence and not law enforcement agency), we should expect similar qualms in many countries if police used similar capabilities. Nonetheless, it seems likely most jurisdictions might allow police to use some data analytics capabilities. Legal systems that protect privacy distinguish between public and private information from unnecessary government surveillance and protect private information. What remains is a socio-political settlement as to what constitutes public and private information.

Conclusion

This article has probably inspired in its readers more questions than it has answers, and it should. No matter how familiar the internet seems in our daily lives, many of its effects remain poorly understood and scarcely explained, especially within broader intellectual and professional communities. Not every potential effect of the internet merits an independent explanation by experts within each field. Perhaps the internet has not sufficiently changed organized crime, for

example, that there should be an entire study examining the internet's effects on the mob. The internet has probably had enough effect on organized crime that scholars should account for it.

This article remedies one obstacle to better understanding the misperception of online information as monolithic. For many people, all information online is the same because it all comes from computers. For police science, however, there are important distinctions between online information sources. The legal and technological remedies available for each information source are different.

Knowing where meaningful, observable boundaries exist groups appropriately similar phenomena together. This article shows why Russian agents searching Hillary Clinton's public websites for emails they can use for spear-phishing is similar to those same agents collecting information on voters from public registrars even though they had disparate effects.[80] Though used for similar effect, hacking into the Democratic National Committee and spear-phishing John Podesta were not similar information sources.[81] Differentiating among the different information sources will make future research more reliable and useful.

This article has also proposed police and police science can use different information sources, sometimes in ways similar to the ways criminals use online information. Criminologists should understand the power of the internet and its information from criminology and policing. Treating the internet as a boogeyman, or at best a useful venue for non-policing activities, short-changes police of a useful tool. By learning the internet's power and limitations in crime prevention, police and criminologists will also learn the internet's power and limitations for crime.

Perhaps the most important question this article only pays glancing attention to is the appropriate role for police online. One need not be a civil libertarian to worry that police may abuse their real-world authority online. Although we may normatively ask the proper role of police online, practical questions should inform future discussions. Only when we know what trade-offs we make by allowing police to engage in certain practices online can we make reasonable policy conclusions.

To the police scientist, the first question they should ask is: of the things that I study, what might online information change, and how? There is no reason to assume online information will change anything at all, but until scholars ask the question we cannot know. The first step in that inquiry is to identify the role information plays in different facets of police science. Having identified information's role, one or more of the information sources discussed in this article may recommend itself for further inquiry. Good inquiries are likely to set off dialogues among the various intellectual stakeholders, but in the dialogue we can progress closer to a full understanding of this new world we live in.

Author

David C. Benson, PhD
Professor of Security and Strategic Studies, School of Advanced Air and Space Studies
david.benson.13@au.af.edu
https://www.airuniversity.af.edu/Portals/10/SAASS/Faculty/DavidBenson_Bio.pdf

Dr. Benson is a Professor of Security and Strategic Studies at the School of Advanced Air and Space Studies. His research focuses on the political effects new information technologies are having on international relations. He served in the US Army, including a deployment to Iraq where he worked as a liaison with local police. He has also supported various NATO Centers of Excellence for counterterrorism.

*Corrections included addressing typographical errors and formatting issues.

Endnotes

[1] Joanne Berman et al., *Twitter Investigation Report* (Albany, NY: Department of Financial Services, 2020), https://www.dfs.ny.gov/Twitter_Report.
[2] Dirk Bergemann and Alessandro Bonatti, "Targeting in Advertising Markets: Implications for Offline Versus Online Media," *The RAND Journal of Economics* 42, no. 3 (2011): 417–43, http://www.jstor.org/stable/23046807.
[3] Stephanie Aceves, "How to Guard Against Cybersecurity Hopelessness," *SC Magazine*, 3 November 3 2021, https://www.scmagazine.com/perspective/policy/how-to-guard-against-cybersecurity-hopelessness; Thomas Rid, *Cyber War Will Not Take Place* (Oxford: Oxford University Press, 2013).
[4] Robert Barr Smith, *Last Hurrah of the James-Younger Gang* (Norman, OK: University of Oklahoma Press, 2001).
[5] David C. Benson, "Why the Internet Is Not Increasing Terrorism," *Security Studies* 23, no. 2 (3 April 2014): 293–328, https://doi.org/10.1080/09636412.2014.905353.
[6] The TCP/IP requires receipt acknowledgement, but some online protocols like UDP do not. *See* IBM, "TCP/IP TCP, UDP, and IP Protocols," IBM z/OS Documentation, last updated 21 March 2022, https://prod.ibmdocs-production-dal-6099123ce774e592a519d7c33db8265e-0000.us-south.containers.appdomain.cloud/docs/en/zos/2.2.0?topic=internets-tcpip-tcp-udp-ip-protocols.
[7] Daniel Gomes, João Miranda, and Miguel Costa, "A Survey on Web Archiving Initiatives," in *Research and Advanced Technology for Digital Libraries*, ed. Stefan Gradmann et al., vol. 6966, Lecture Notes in Computer Science (Berlin, Heidelberg: Springer Berlin Heidelberg, 2011), 408–20, https://doi.org/10.1007/978-3-642-24469-8_41.
[8] Johannes Haupt et al., "Robust Identification of Email Tracking: A Machine Learning Approach," *European Journal of Operational Research* 271, no. 1 (11 June 2018): 341–56, https://doi.org/10.1016/j.ejor.2018.05.018.
[9] Wikileaks, *The Wikileaks Files: The World According to US Empire* (Brooklyn, NY: Verso Books, 2016).
[10] Zhanna Malekos Smith and Eugenia Lostri, *The Hidden Costs of Cybercrime* (San Jose, CA: McAfee, 2021).
[11] Steve Morgan, "Special Report: Cyberwarfare In The C-Suite," *Cybercime Magazine*, 13 November 2020, https://cybersecurityventures.com/cybercrime-damages-6-trillion-by-2021/.
[12] Madeline Fox, "Ransomware, Phishing and Cyberattacks Are Increasingly Hitting Wisconsin School Districts, Most Recently in Janesville," *Wisconsin Public Radio*, 28 October 2021, https://www.wpr.org/ransomware-phishing-and-cyberattacks-are-increasingly-hitting-wisconsin-school-districts-most; Titilope Joseph, "N1.87b Fraud: SFU Arrest Alleged Suspect Who Hacks Into Banks Server," *Independent Newspaper Nigeria*, 12 August 2021, https://independent.ng/n1-87b-fraud-sfu-arrest-alleged-suspect-who-hacks-into-banks-server/; Sarah Robson, "Cyber Attack Continues to Cause Delays in Waikato DHB," *RNZ*, 23 August 2021, https://www.rnz.co.nz/news/national/449756/cyber-attack-continues-to-cause-delays-in-waikato-dhb.
[13] "Cyberstalking: Two Federal Cases Illustrate the Consequences of Sextortion," News, FBI, 30 October 2018, https://www.fbi.gov/news/stories/sentences-in-separate-cyberstalking-cases-103018.
[14] Michael Martelle and Rosemary Tropeano, "Tainted Trove," National Security Archive, The George Washington University, 20 February 2019, https://nsarchive.gwu.edu/news/cyber-vault/2019-02-20/tainted-trove.
[15] Andrew Eich, "Chinese APT Hafnium Attacking Microsoft Exchange Servers," Cybersecurity, University of Hawai'i-West Oahu, 25 March 2021, https://westoahu.hawaii.edu/cyber/uncategorized/chinese-apt-hafnium-attacking-microsoft-exchange-servers/.
[16] Team Atlas, "SolarWinds: Advancing the Story," Threat Intel Portal, RiskIQ, accessed June 2022, https://community.riskiq.com/article/9a515637.
[17] Office of Public Affairs, "Two Iranian Nationals Charged for Cyber-Enabled Disinformation and Threat Campaign Designed to Influence the 2020 U.S. Presidential Election," Department of Justice, 18 November 2021, https://www.justice.gov/opa/pr/two-iranian-nationals-charged-cyber-enabled-disinformation-and-threat-campaign-designed.
[18] Bellingcat Investigation Team, "Berlin Assassination: New Evidence on Suspected FSB Hitman Passed to German Investigators," Bellingcat, 19 March 2021, https://www.bellingcat.com/news/2021/03/19/berlin-assassination-new-evidence-on-suspected-fsb-hitman-passed-to-german-investigators/.
[19] Sebastian Anthony, "The first rule of zero-days is no one talks about zero-days (so we'll explain)," *Ars Technica*, 20 October 2015, http://arstechnica.co.uk/security/2015/10/the-rise-of-the-zero-day-market/.
[20] Lillian Ablon, Martin C. Libicki, and Andrea A. Golay, *Markets for Cybercrime Tools and Stolen Data: Hackers' Bazaar* (Santa Monica, CA: RAND Corporation, 2014).
[21] Keman Huang, Michael Siegel, and Stuart Madnick, "Cybercrime-as-a-Service: Identifying Control Points to Disrupt," Working Paper CISL# 2017-17 (Massachusetts Institute of Technology, 2017), http://web.mit.edu/smadnick/www/wp/2017-17.pdf.

[22] Huang, Siegel, and Madnick, "Cybercrime-as-a-Service"; Steven Melendez, "Hackers On Demand," *Fast Company*, 29 May 2015, http://www.fastcompany.com/3043016/the-cybercrime-service-economy.
[23] Kristen Korosec, "The 25 Most Common Passwords of 2017 Include 'Star Wars'," *Fortune*, 19 December 2017, http://fortune.com/2017/12/19/the-25-most-used-hackable-passwords-2017-star-wars-freedom/.
[24] Katie Petrillo, "New Research: Psychology of Passwords, Neglect is Helping Hackers Win," LastPass, 1 May 2018, https://blog.lastpass.com/2018/05/psychology-of-passwords-neglect-is-helping-hackers-win/.
[25] Olabode Anise and Kyle Lady, "State of the Auth: Experiences and Perceptions of Multi-Factor Authentication," Duo Labs, 7 November 2017, https://duo.com/blog/state-of-the-auth-experiences-and-perceptions-of-multi-factor-authentication; Iain Thomson, "Who's Using 2fa? Sweet FA. Less Than 10% of Gmail Users Enable Two-Factor Authentication," *The Register*, 17 January 2018, https://www.theregister.co.uk/2018/01/17/no_one_uses_two_factor_authentication/; Brian Krebs, "State Govts. Warned of Malware-Laden CD Sent Via Snail Mail from China," Krebs on Security, 27 July 2018, https://krebsonsecurity.com/2018/07/state-govts-warned-of-malware-laden-cd-sent-via-snail-mail-from-china/.
[26] Jeff White, "Analyzing Attacks Against Microsoft Exchange Server with China Chopper Webshells," Unit 42, 8 March 2021, https://unit42.paloaltonetworks.com/china-chopper-webshell/.
[27] Christiaan Beek, Cedric Cochin, and Raj Samani, "Additional Analysis into the SUNBURST Backdoor," McAfee, 17 December 2020, https://www.mcafee.com/blogs/other-blogs/mcafee-labs/additional-analysis-into-the-sunburst-backdoor/; Jon R. Lindsay, "Stuxnet and the Limits of Cyber Warfare," *Security Studies* 22, no. 3 (2013): 365–404, https://doi.org/10.1080/09636412.2013.816122.
[28] "Highly Evasive Attacker Leverages SolarWinds Supply Chain to Compromise Multiple Global Victims With SUNBURST Backdoor," Threat Research, FireEye, 13 December 2020, https://www.fireeye.com/blog/threat-research/2020/12/evasive-attacker-leverages-solarwinds-supply-chain-compromises-with-sunburst-backdoor.html.
[29] Lawrence Abrams, "REvil ransomware hits Managed.com hosting provider, 500K ransom," BleepingComputer, 18 November 2020, https://www.bleepingcomputer.com/news/security/revil-ransomware-hits-managedcom-hosting-provider-500k-ransom/.
[30] Tara Seals, "Las Vegas Students' Personal Data Leaked, Post-Ransomware Attack," *Threatpost*, 29 September 2020, https://threatpost.com/las-vegas-students-data-leaked-ransomware/159645/.
[31] Steve Alder, "Lawsuit Alleges Ransomware Attack Resulted in Hospital Baby Death," *HIPAA Journal*, 4 October 2021, https://www.hipaajournal.com/lawsuit-alleges-ransomware-attack-resulted-in-hospital-baby-death/; Dan Goodin, "A Patient Dies After a Ransomware Attack Hits a Hospital," *Wired*, 19 September 2020, https://www.wired.com/story/a-patient-dies-after-a-ransomware-attack-hits-a-hospital/.
[32] ThreatConnect Research Team, "Shiny Object? Guccifer 2.0 and the DNC Breach," ThreatConnect, June 2016, https://threatconnect.com/guccifer-2-0-dnc-breach/.
[33] Elizabeth Blosfield, "Ransomware Has Been a 'Game Changer' for Cyber Insurance," *Insurance Journal*, 30 August 2021, https://www.insurancejournal.com/news/national/2021/08/30/628672.htm.
[34] Collin Eaton and Dustin Volz, "Colonial Pipeline CEO Tells Why He Paid Hackers a $4.4 Million Ransom," *Wall Street Journal*, 19 May 2021, https://www.wsj.com/articles/colonial-pipeline-ceo-tells-why-he-paid-hackers-a-4-4-million-ransom-11621435636.
[35] Martin Roesler, "What We Can Learn From the Bangladesh Central Bank Cyber Heist," Trend Micro, 16 March 2016, https://blog.trendmicro.com/trendlabs-security-intelligence/what-we-can-learn-from-the-bangladesh-central-bank-cyber-heist/.
[36] Brett Hawkins, "Case Study: The Home Depot Data Breach," SANS Institute, 27 October 2015, https://www.sans.org/white-papers/36367/; Rachel Weiner, "Hacker Linked to Target Data Breach Gets 14 Years in Prison," *Washington Post*, 21 September 2018, https://www.washingtonpost.com/local/public-safety/hacker-linked-to-target-data-breach-gets-14-years-in-prison/2018/09/21/839fd6b0-bd17-11e8-b7d2-0773aa1e33da_story.html.
[37] James Scott, "ICIT Analysis - Equifax: America's In-Credible Insecurity Part One," Institute for Critical Infrastructure Technology, 7 September 2017, https://icitech.org/icit-analysis-equifax-americas-in-credible-insecurity-part-one/; James Scott, "Equifax: The Hazards of Dragnet Surveillance Capitalism Part 2: Just Another Data Breach? Or C-Suite Criminal Negligence?," Institute for Critical Infrastructure Technology, October 2017, https://icitech.org/equifax-the-hazards-of-dragnet-surveillance-capitalism-part-2-just-another-data-breach-or-c-suite-criminal-negligence/.
[38] Ellen Nakashima and Reed Albergotti, "An Australian Hacking Firm Solved FBI's iPhone Problem," *The Washington Post*, 15 April 2021, http://www.proquest.com/usnews/docview/2512591903/citation/E0A23FBF5C5C4C65PQ/3.
[39] "FBI's Encrypted Phone Platform Infiltrated Hundreds of Criminal Syndicates; Result Is Massive Worldwide Takedown," U.S. Attorney's Office, Southern District of California, 8 June 2021, https://www.justice.gov/usao-sdca/pr/fbi-s-encrypted-phone-platform-infiltrated-hundreds-criminal-syndicates-result-massive.
[40] Asheer Malhotra, "IndigoDrop Spreads via Military-Themed Lures to Deliver Cobalt Strike," Cisco Talos Intelligence Blog, 22 June 2020, http://blog.talosintelligence.com/2020/06/indigodrop-maldocs-cobalt-strike.html.

[41] Bill Marczak et al., "Missing Link: Tibetan Groups Targeted with 1-Click Mobile Exploits," Targeted Threats, The Citizen Lab, 24 September 2019, https://citizenlab.ca/2019/09/poison-carp-tibetan-groups-targeted-with-1-click-mobile-exploits/.
[42] M. Adam Mahmood et al., "Moving Toward Black Hat Research in Information Systems Security: An Editorial Introduction to the Special Issue," *MIS Quarterly* 34, no. 3 (2010): 431–33, https://doi.org/10.2307/25750685.
[43] Riana Pfefferkorn, "America's Anti-Hacking Laws Pose a Risk to National Security," *TechStream*, The Brookings Institution, 7 September 2021, https://www.brookings.edu/techstream/americas-anti-hacking-laws-pose-a-risk-to-national-security/.
[44] Curtis Peterson, "23 Social Engineering Attacks You Need to Shut Down," *Smartfile*, 30 June 2020, https://www.smartfile.com/blog/social-engineering-attacks/.
[45] Nex, "Operation Kingphish: Uncovering a Campaign of Cyber Attacks Against Civil Society in Qatar and Nepal," *Amnesty Insights,* Amnesty International, 14 February 2017, https://medium.com/amnesty-insights/operation-kingphish-uncovering-a-campaign-of-cyber-attacks-against-civil-society-in-qatar-and-aa40c9e08852; Fabien Goa and James Lynch, "Beyond Fake News: An Investigation into the Murky World of Fake Campaigns," *Amnesty Global Insights*, Amnesty International, 21 December 2016, https://medium.com/amnesty-insights/beyond-fake-news-an-investigation-into-the-murky-world-of-fake-campaigns-f4af8118844b.
[46] "Insider Threat – Cyber," Cybersecurity & Infrastructure Security Agency, https://www.cisa.gov/insider-threat-cyber.
[47] "FBI Arrests Glendale Man In 'Sextortion' Case," U.S. Attorney's Office, Central District of California, 29 January 2013, https://www.justice.gov/usao-cdca/pr/fbi-arrests-glendale-man-sextortion-case.
[48] Ellen McCarthy, "What Is Catfishing? A Brief (and Sordid) History." *Washington Post*, 9 January 2016, https://www.washingtonpost.com/news/arts-and-entertainment/wp/2016/01/09/what-is-catfishing-a-brief-and-sordid-history/?utm_term=.c6cbd9ce84a4; Phillip Knightly, "The History of the Honey Trap," *Foreign Policy*, 12 March 2010, https://foreignpolicy.com/2010/03/12/the-history-of-the-honey-trap/.
[49] "Cyber Actors Use Online Dating Sites to Conduct Confidence/Romance Fraud and Recruit Money Mules," Public Service Announcement, Federal Bureau of Investigation, 5 August 2019, https://www.ic3.gov/Media/Y2019/PSA190805.
[50] Pat Engebretson, *The Basics of Hacking and Penetration Testing: Ethical Hacking and Penetration Testing Made Easy*, 2nd ed. (Amsterdam, Boston: Syngress, 2013), 127–40.
[51] Brian Fung, "Why This Google Docs Phishing Attack Is Particularly Sneaky," *Washington Post*, 3 May 2017, https://www.washingtonpost.com/news/the-switch/wp/2017/05/03/why-this-google-docs-phishing-attack-is-particularly-sneaky/?noredirect=on&utm_term=.3c9f2d3c2404.
[52] Holger Schulze, *Insider Threat Report* (Cybersecurity Insiders, 2020), https://www.cybersecurity-insiders.com/wp-content/uploads/2019/11/2020-Insider-Threat-Report-Gurucul.pdf.
[53] "Russian Pleads Guilty to Tesla Ransomware Plot," *BBC News*, 20 March 2021, https://www.bbc.com/news/world-us-canada-56469475.
[54] Wikileaks, *The Wikileaks Files.*
[55] Andy Greenberg, "The Attack That Broke Twitter Is Hitting Dozens of Companies," *Wired*, 18 August 2020, https://www.wired.com/story/phone-spear-phishing-twitter-crime-wave/.
[56] Margot Cleveland et al., "Trends in the International Fight Against Bribery and Corruption," *Journal of Business Ethics* 90 (2009): 199–244, https://doi.org/10.1007/s10551-010-0383-7.
[57] Hsinchun Chen, Roger H. L. Chiang, and Veda C. Storey, "Business Intelligence and Analytics: From Big Data to Big Impact," *MIS Quarterly* 36, no. 4 (2012): 1165–88, https://doi.org/10.2307/41703503; Viktor Mayer-Schönberger and Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work, and Think* (Boston: Houghton Mifflin Harcourt, 2013).
[58] Johnathan A. Segal, "Social Media Use in Hiring: Assessing the Risks," HR Magazine, 2018, https://www.shrm.org/hr-today/news/hr-magazine/pages/0914-social-media-hiring.aspx; Saige Driver, "Keep It Clean: Social Media Screenings Gain in Popularity," *Business News Daily*, 3 April 2018, https://www.businessnewsdaily.com/2377-social-media-hiring.html.
[59] Craig Stedman and Rachel Lebeaux, "What Is Big Data as a Service (BDaaS)?" Search CIO, Tech Target, June 2021, https://searchcio.techtarget.com/definition/big-data-as-a-service-bdaas.
[60] *See* Twint, https://github.com/twintproject/twint.
[61] *See* Creepy by Ilekrojohn, http://ilektrojohn.github.com/creepy/.
[62] Carter Jernigan and Behram F. T. Mistree, "Gaydar: Facebook Friendships Expose Sexual Orientation," *First Monday* 14, no. 10, 2009, https://doi.org/10.5210/fm.v14i10.2611.
[63] Mauricio Santillana et al., "Using Clinicians' Search Query Data to Monitor Influenza Epidemics," *Clinical Infectious Diseases* 59, no. 10 (2014): 1446–50, http://www.jstor.org/stable/24032319.
[64] Nathan Kallus, "Predicting Crowd Behavior with Big Public Data," in *Proceedings of the 23rd International Conference on World Wide Web* (Seoul, Korea: ACM, 2014), 625–30, https://doi.org/10.1145/2567948.2579233.

[65] F. A. Hayek, "The Use of Knowledge in Society," *The American Economic Review* 35, no. 4 (1945): 519–30, http://www.jstor.org/stable/1809376.
[66] Robin Young and Serena McMahon, "Kinsa Smart Thermometer Data Predicts New COVID-19 Spikes Weeks Before CDC," *WBUR*, 25 June 2020, https://www.wbur.org/hereandnow/2020/06/25/covid-19-kinsa-smart-thermometer.
[67] Kai-Fu Lee, *AI Superpowers: China, Silicon Valley, and the New World Order* (New York: Houghton Mifflin Harcourt, 2018).
[68] Chen, Chiang, and Storey, "Business Intelligence and Analytics"; Michel Wedel and P.K. Kannan, "Marketing Analytics for Data-Rich Environments," *Journal of Marketing* 80, no. 6 (2016): 97–121, http://www.jstor.org/stable/44134975.
[69] Ellen Hughes-Cromwick and Julia Coronado, "The Value of US Government Data to US Business Decisions," *The Journal of Economic Perspectives* 33, no. 1 (2019): 131–46, http://www.jstor.org/stable/26566980.
[70] "Twitter API," Developer Platform, Twitter, accessed 4 October 2021, https://developer.twitter.com/en/docs/twitter-api.
[71] Andrea Valdez, "Everything You Need to Know about Facebook and Cambridge Analytica," *Wired*, 23 March 2018, https://www.wired.com/story/wired-facebook-cambridge-analytica-coverage/.
[72] "Turn Websites into Data," Scraping Hub, 2018, https://scrapinghub.com/.
[73] "General Data Protection Regulation," Pub. L. No. 119/1, 2016 OJ (2016).
[74] Raphael Satter, Jeff Donn, and Chad Day, "Inside Story: How Russians Hacked the Democrats' Emails," *Associated Press News*, 4 November 2017, https://www.apnews.com/dea73efc01594839957c3c9a6c962b8a.
[75] Ed Adamczyk, "Guccifer Pleads Guilty to Hacking Emails of Clinton, Ex-Presidents," *UPI*, 25 May 2016, https://www.upi.com/Top_News/US/2016/05/25/Hacker-Guccifer-pleads-guilty-to-email-break-ins/9481464195286/.
[76] Peter Funt, "Burglars Are Following You on Facebook," *Wall Street Journal*, 11 September 2019, https://www.wsj.com/articles/burglars-are-following-you-on-facebook-11568244205.
[77] "Fraud Alert: COVID-19 Scams," Office of Inspector General, U.S. Department of Health and Human Services, 24 December 2020, https://oig.hhs.gov/fraud/consumer-alerts/fraud-alert-covid-19-scams/.
[78] Sven Probst, Nic Carrington, and Andra Horwat, "COVID-19 – A Backdoor to Increased Fraud Risk?" (Deloitte Switzerland, 2021), https://www2.deloitte.com/ch/en/pages/financial-advisory/articles/covid-19-operating-in-the-new-normal-fraud-risk.html.
[79] David P. Fidler and Sumit Ganguly, eds., *The Snowden Reader* (Bloomington, Indiana: Indiana University Press, 2015).
[80] Craig Mauger, "Michigan Voter Information Wasn't Hacked, Benson's Office Says," *The Detroit News*, 1 September 2020, https://www.detroitnews.com/story/news/politics/2020/09/01/michigan-voter-information-wasnt-hacked-secretary-state-benson-says/5682092002/; Gregory Krieg and Tal Kopan, "Is This the Email That Hacked John Podesta's Account?" *CNN*, 30 October 2016, https://www.cnn.com/2016/10/28/politics/phishing-email-hack-john-podesta-hillary-clinton-wikileaks/index.html.
[81] Satter, Donn, and Day, "Inside Story."