# THE GRAY DATA ZONE:
# OSINT, ETHICS, AND THE LEGAL BOUNDARIES OF DIGITAL INVESTIGATIONS

By **Jody O'Guinn, MPA, CHS-III, CCFI, WVTS**

Loss prevention (LP) professionals operate in an era where the volume of available data is immense, and the speed of access is nearly instantaneous. Yet, the ease of acquiring information is counterbalanced by a growing complexity in using it responsibly. Private investigators, often working in tandem with corporate LP teams, navigate what can best be described as the "Gray Data Zone," or the ambiguous territory between what is public and what is permissible, what can be accessed and what can be used, and what is informative but not necessarily admissible. It is within this gray zone that open-source intelligence, or OSINT, truly operates.

If your organization has ever asked a contractor to "check social," hired a third-party data broker, or gathered online intelligence to support an investigation, you are already working in this zone. Understanding where that line lies, and how to stay on the right

Who is Danny / Shutterstock.com

side of it, is now critical not only to the integrity of investigations but to the long-term credibility of the LP profession itself.

## What OSINT Really Is—and Why It Matters

OSINT, in its purest definition, is the collection and analysis of information that is lawfully available to the public. It extends far beyond the surface of the internet. It includes open websites, social platforms, government records, digital marketplaces, and even the metadata embedded in images and videos. Investigators use these sources to uncover behavioral patterns, link associates, corroborate stories, and identify anomalies that point toward theft, fraud, or misconduct.

For LP teams, OSINT has become indispensable. It allows investigators to trace fencing networks, monitor employee activity that may signal internal theft, and identify organized retail crime operations that span multiple states. It helps confirm addresses, associate accounts with specific identities, and anticipate where stolen goods might resurface. When properly applied, it saves time, reduces physical surveillance costs, and focuses investigative resources where they will yield the greatest results.

However, the same tools that make OSINT powerful also make it perilous. The assumption that "if it's online, it's fair game" is a misconception that has led many professionals into ethical and legal jeopardy. The difference between effective intelligence gathering and an unlawful invasion of privacy often hinges on how data is collected and subsequently disseminated, not just what data is collected.

## Where the Gray Begins: Law, Policy, and Platform Rules

The first misconception many investigators fall prey to is equating visibility with free use. Just because a social media post or photo is publicly viewable does not mean it can be downloaded, republished, or scraped in violation of a platform's terms of service. These terms carry contractual weight, and violations can result in account suspension and legal liability. Several court cases have already addressed whether automated collection of publicly available data, known as scraping, violates computer misuse statutes. The outcomes vary, but the trend is clear: regulators are increasingly willing to draw limits around how far "open source" can stretch.

Another critical issue is the concept of privacy expectation. Even if information is accessible online, courts may still recognize a reasonable expectation of privacy depending on the context. For instance, collecting data from restricted groups, closed forums, or private chat rooms using deception can cross ethical and legal lines. Investigators must ask not only whether they can access information, but how that access was achieved and whether it would withstand scrutiny in court.

LP professionals must also be mindful of sector-specific laws. The Fair Credit Reporting Act (FCRA) applies whenever collected information is used to make employment decisions, such as hiring, discipline, or termination. A company that uses OSINT in an employee investigation must ensure that consent, dispute rights, and adverse-action protocols are followed. Similarly, privacy frameworks like the California Consumer Privacy Act (CCPA) or Europe's General Data Protection Regulation (GDPR) impose obligations regarding how data is stored, shared, and deleted. In other words, the LP function cannot hide behind the "we're not law enforcement" defense. Compliance applies to anyone processing personal data, regardless of intent.

**LP PROFESSIONALS MUST ALSO BE MINDFUL OF SECTOR-SPECIFIC LAWS. THE FAIR CREDIT REPORTING ACT (FCRA) APPLIES WHENEVER COLLECTED INFORMATION IS USED TO MAKE EMPLOYMENT DECISIONS, SUCH AS HIRING, DISCIPLINE, OR TERMINATION.**

Because of these complexities, OSINT collection should be guided by legal counsel whenever possible. When investigations are conducted under the direction of an attorney, privilege protections may apply, and documentation can be structured to reflect a legitimate investigative purpose. This does not shield misconduct, but it does ensure that the investigation follows a defensible framework from start to finish.

## The Ethical Dimension: Doing It Right, Not Just Doing It Legally

If the law determines what you may do, ethics determine what you should do. Ethics are not an afterthought; they are a strategic safeguard. A reputation for fairness and restraint can determine whether a prosecutor trusts your report, whether a jury believes your findings, and whether your organization retains public credibility after an investigation concludes.

An ethical approach to OSINT begins with legitimacy and necessity. Investigators should collect information only when it serves a legitimate business purpose tied directly to a documented risk or incident, such as fraud, workplace violence, or theft. The scope should be narrowly defined to avoid unnecessary intrusion. Collecting "everything available" is not only inefficient; it may expose the company to privacy claims or data breach liability if that information is mishandled.

Equally important is truthfulness. Deceptive identities, pretext accounts, or impersonation can sometimes seem like shortcuts to information, but they are rarely justified and often illegal. If investigators must operate under an assumed identity, for example, when verifying the sale of stolen goods online, there must be explicit written authorization, a clearly defined scope, and a documented justification for why it was necessary.

Finally, data security and accountability close the ethical loop. Investigators are custodians of sensitive information, and that responsibility extends beyond the case's conclusion. Proper retention schedules, secure storage, and routine audits ensure that information is used ethically and protected against misuse.

## The Lead vs. Evidence Divide

One of the most important distinctions in OSINT work is the line between information that leads to an investigative direction and information that can stand as evidence. Too often, investigators treat screenshots or online posts as conclusive

proof, only to see them excluded because the method of collection compromised authenticity. OSINT should first be viewed as a tool for lead development. Once a digital clue points to a viable path, the investigator should shift to formal evidentiary collection methods that preserve metadata, establish a chain of custody, and comply with evidentiary standards.

When an investigator captures a social media post or marketplace listing, it should be immediately preserved with a time stamp, URL, and cryptographic hash value to ensure integrity. Tools designed for forensic web capture—free tools such as Magnet WPS, Forensic OSINT, and PageRecon—can create admissible records, while simple screenshots, especially without context, invite doubt. The difference between a useful lead and admissible evidence can be the difference between an arrest and a dismissed case.

## The Reliability Challenge

In the era of disinformation, assessing credibility is as important as finding data itself. Online information must be weighed on a spectrum of reliability. A self-asserted claim from an anonymous user, for example, carries little evidentiary weight compared to an official record verified by an external source. Cross-platform corroboration, such as matching profile photos, handles, or geolocation tags, adds confidence, but investigators should never assume that online identities are consistent or truthful.

One practical approach is to rate each data point by reliability, explaining why it should be considered strong or weak. Including such assessments in investigative reports demonstrates analytical rigor and helps decision makers understand the context behind the findings. This discipline not only improves case quality but also protects investigators from the accusation of bias or selective reporting.

## The Undercover Temptation

Operating undercover in the digital realm can be effective but dangerous. Using "sock puppet" accounts, false personas created to observe or engage with suspects, requires clear boundaries. If investigators infiltrate closed groups or pretend to be someone they are not, they may violate both platform rules and, in some jurisdictions, laws against deceptive access and practices. The line between legitimate observation and entrapment can blur quickly.

Organizations should maintain written policies governing the use of assumed identities, defining who can authorize them, how credentials are managed, and under what circumstances they may be used. Investigators should never impersonate real individuals or falsify professional credentials. Even when done lawfully, such tactics should be employed only when absolutely necessary and documented meticulously.

**AS TECHNOLOGY EVOLVES, INVESTIGATORS ARE INCREASINGLY TEMPTED TO AUTOMATE DATA COLLECTION THROUGH SCRAPING TOOLS AND BOTS. WHILE AUTOMATION CAN SAVE TIME, IT ALSO MULTIPLIES RISK.**

## Automation, Scraping, and the Allure of Scale

As technology evolves, investigators are increasingly tempted to automate data collection through scraping tools and bots. While automation can save time, it also multiplies risk. Automated scraping that circumvents access controls or overwhelms a platform's servers may constitute unauthorized access. Furthermore, automated tools can inadvertently capture personal or sensitive data that exceeds the investigative scope, creating privacy liabilities.

When possible, manual collection is safer for critical evidence. It ensures human oversight, captures context, and avoids the perception of indiscriminate data harvesting. If automation must be used, it should be transparent, rate-limited, and compliant with platform terms. Every data collection effort should be documented, including how, when, and by whom the data was retrieved, so that the process can withstand legal or regulatory review.

## The Double-Edged Sword of Social Media

Social media is perhaps the richest and riskiest OSINT source for LP. It has exposed countless theft rings, insider misconduct cases, and fraud networks, but it is also rife with pitfalls. Misidentification is common; people share names, photos are repurposed, and content may be years old or altered. Before drawing conclusions from social media data, investigators should verify time stamps, confirm identities, and preserve full-page captures that include the context of posts.

Private messages and direct communications, while tempting, are generally off-limits without legal authority. Attempting to access them by deception or without consent can expose investigators to significant liability. When a case depends on private messages or account records, law enforcement should be engaged early to pursue subpoenas or warrants.

Even public content must be preserved carefully. Screenshots alone are insufficient; forensic web capture tools can record the underlying metadata, helping authenticate posts in court. The key is discipline: capture the evidence lawfully, preserve it thoroughly, and let the legal process do the rest.

## Building Cases Ethically: Organized Retail Crime

Organized retail crime cases often rely on link analysis, the mapping of relationships between people, places, vehicles, and online accounts. OSINT can be invaluable here, but it must be handled with precision. The best investigators start by defining known anchors, such as case numbers, SKU identifiers, or confirmed suspects. They then map relationships between social media profiles, associates, addresses, and online marketplaces. Each link is supported with evidence, time-stamped, and rated for reliability.

The objective is not to collect everything but to build enough confidence in a small set of connections to justify further action. Once a network is reasonably established, investigators can coordinate with law enforcement to validate findings through lawful process, such as subpoenas for marketplace records or payment data. In doing so, they maintain a clear boundary between private sector intelligence and public sector enforcement, ensuring the integrity of both.

# Stop theft before it starts...



SONR™ Pusher

SONR™ Spiral Anti-Sweep Hook™


LM Tag™ Loop with SONR™

LM Tag™ with SONR™

## SONR™ System - Smarter, Stronger Store Security

Protect your products where it matters most, right at the shelf. SONR™ is a complete wireless ecosystem that detects product activity and instantly alerts store staff.

From hooks and shelf pushers to LM Tags and more, SONR™ solutions deter theft by letting offenders know items are being monitored, all while relaying activity to the SONR Echo Box up to 200 feet away.

Easy to install and scale to any category or store, SONR™ integrates seamlessly with your existing security systems to create a powerful tiered alert network. The result? Immediate deterrence, real-time awareness, and smarter protection for every product.

SONR™ Echo Box

Call us at **800.422.2547** or visit **www.siffron.com**

siffron®
always on

## Collaboration with Law Enforcement

Partnerships between corporate LP teams, private investigators, and law enforcement are essential in combating theft and fraud that cross jurisdictions. Yet these relationships must be built on mutual respect for legal process. Private investigators cannot issue subpoenas or compel records, and they should never attempt to act as de facto police. Their role is to preserve what is publicly visible, package it coherently, and provide law enforcement with a concise, well documented summary of findings.

Law enforcement agencies appreciate organized, authenticated evidence packets far more than massive data dumps. A well-prepared packet includes a clear methodology, an index of artifacts, time-stamped captures, and a narrative that explains relevance. It should also maintain chain-of-custody documentation for each artifact. When investigators work within these parameters, they not only enhance prosecutorial success but also reinforce the credibility of the private sector as a responsible partner in public safety.

## Documentation and Governance

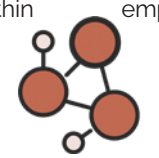
A defensible OSINT program begins with governance. Every organization conducting online investigations should have a written playbook outlining purpose, roles, collection standards, review processes, and retention protocols. The playbook should specify who is authorized to conduct investigations, under what circumstances, and what tools or vendors may be used.

Collection standards should emphasize manual capture for key artifacts, strict adherence to privacy laws, and explicit bans on illegal access or pretexting. Retention schedules must be defined, and non-responsive or irrelevant data must be deleted in a timely manner. Vendors should be contractually obligated to comply with the same standards, disclose their methods, and allow audits of their processes. Regular internal audits and annual training sessions reinforce these expectations and ensure that ethical practice becomes institutional habit rather than individual discretion.

## The Human Element: Bias and Fairness

Perhaps the most overlooked risk in OSINT is human bias. Analysts bring assumptions and perspectives to their work, and those assumptions can color how data is interpreted. Confirmation bias—seeing what you expect to see—can lead investigators to misread neutral content as incriminating. To counter this, investigators must deliberately seek disconfirming evidence, document alternative explanations, and have a second analyst review key findings.

Fairness also extends to how subjects are treated. When investigations involve employees, contractors, or customers, the principle of proportionality should guide action. Collect only what is necessary, avoid collateral exposure of uninvolved individuals, and redact or anonymize unrelated data when sharing internally. Ethical restraint is not weakness; it is the mark of professionalism.



**LP THRIVES ON SPEED, ACTING QUICKLY TO STEM LOSSES, PROTECTING EMPLOYEES, AND PREVENTING HARM. OSINT ENABLES THAT SPEED, BUT UNRESTRAINED USE INVITES REGRET. THE GRAY DATA ZONE WILL ALWAYS EXIST; TECHNOLOGY EVOLVES FASTER THAN LAW.**

## Case Illustrations

In one recent case, an LP team investigating excessive returns discovered that the same customer was selling identical merchandise online. OSINT revealed the reseller's profile, which included photos of the exact products with matching SKU labels. Investigators carefully captured the listings with time stamps and hash values before engaging law enforcement. The resulting warrants led to the dismantling of a regional fencing operation. The success hinged not on luck, but on disciplined, ethical OSINT work.

In another case, an employee posted photos from a restricted stockroom on a private social media group. Rather than overreaching by demanding access to the entire group, investigators focused narrowly on the images provided by a tipster, verified time stamps against security logs, and conducted a structured interview. The employee eventually admitted to the violation, and the company took appropriate corrective action without overstepping legal boundaries.

These examples highlight the difference between precision and overreach. Ethical OSINT is about restraint as much as discovery.

## The Path Forward: Speed with Integrity

LP thrives on speed, acting quickly to stem losses, protecting employees, and preventing harm. OSINT enables that speed, but unrestrained use invites regret. The Gray Data Zone will always exist; technology evolves faster than law. The challenge for LP professionals and private investigators is to build systems that adapt responsibly.

That means treating OSINT not as a shortcut, but as a discipline grounded in legality, ethics, and documentation. It means training teams to differentiate between what they *can* find and what they *should* use. It also means recognizing that credibility is a form of capital. When your investigations are fair, your findings are trusted.

The ultimate goal is balance, harnessing the immense power of open data without crossing into intrusion or impropriety. In the end, how we investigate says as much about our professionalism as what we uncover. The public's trust, and the justice that flows from it, depends on our ability to operate in the Gray Data Zone with clarity, integrity, and purpose. ●

**Jody O'Guinn, MPA, CHS-III, CCFI, WVTS,** is a retired police chief with thirty-four years of law enforcement experience and the CEO and principal investigator of Calabash Investigative Consultants, LLC, a Georgia-licensed investigative firm. A veteran SWAT and regional WMD SRT commander, he also held key roles with the DEA and MEGSI narcotics task forces. A graduate of the FBI National Academy and LEEDS program at Quantico, he holds an MPA from Southern Illinois University. He pioneered the world's first board-certified cryptocurrency forensic investigator course and serves as director of law enforcement operations at Baker Group International.