Global Digital Forensics · Guide edition 2026-11-03 · evestigate.com/insights/the-covert-devaluation-of-us-businesses-hackers-stealing-ip/

SOURCE CODE AND IP PROTECTION

Detecting the Exfiltration of Source Code and Research Assets

When a repository is cloned or a design file is exported without authorization, operations carry on as usual and the loss can surface months later as a competitor's product. Lock down the repository audit logs and developer workstations now, under counsel, while the trail is within retention. Call (800) 868-8189 at any hour.

Published June 16, 2014 · Updated October 30, 2026

Primary Reference for This Guide

Start with NIST SP 800-86, Guide to Integrating Forensic Techniques into Incident Response.

By Robert Knudsen, Computer forensics and cybersecurity examiner, Global Digital Forensics

Why Silent Intellectual Property Theft Is More Damaging Than Loud Ransomware

While a ransomware attack announces itself immediately by locking screens and demanding payment, state sponsored actors and unscrupulous competitors stealing intellectual property want you to never know they were there. When an intruder or insider quietly clones a Git source code repository, exports CAD engineering schematics or copies pharmaceutical and chemical research data, your daily operations continue uninterrupted, until a foreign or domestic competitor releases a knockoff product at half your R&D cost, eroding your enterprise valuation.

Combining Network Exfiltration Forensics With Expert Source Code Review

Global Digital Forensics protects technology companies, manufacturers and financial firms through both proactive security hardening and reactive litigation support. Under the direction of intellectual property counsel, our forensic examiners audit Git version control logs, developer workstation artifacts and cloud egress telemetry to trace repository cloning and test whether it was authorized. When litigation advances to inspecting a competitor software product under a protective order, our SourceScan platform and source code expert witnesses dissect millions of lines of code to identify structural copying, copied comments and shared algorithmic bugs even when variable names were disguised.

© 2026 Global Digital Forensics, Inc. All rights reserved.

Copying Versus Independent Development Is the Question the Statutes Ask

Two federal statutes frame nearly every intellectual property exfiltration matter we support, and each one tells the examiner what to look for. The Defend Trade Secrets Act, at 18 U.S.C. 1839, defines improper means to include theft, breach of a duty to maintain secrecy and "espionage through electronic or other means", and then states that improper means "does not include reverse engineering, independent derivation, or any other lawful means of acquisition". The defendant's answer in a source code case is almost always independent derivation, so the examination traces a path as well as a resemblance. The Computer Fraud and Abuse Act, at 18 U.S.C. 1030, reaches a person who intentionally accesses a computer without authorization or exceeds authorized access and obtains information, and after the Supreme Court narrowed exceeds authorized access in 2021, counsel look to the record for whether the actor reached areas of the system beyond their authorized access, as distinct from using permitted access for a bad purpose.

Those definitions shape the two halves of the work. The access half asks how the code or design data left: which account cloned the repository, from what address, with which SSH key or personal access token, and whether that account was permitted to do so. The similarity half asks whether the competitor's product contains the client's code or data, and whether the overlap is explained by shared open source libraries, common frameworks or published specifications rather than copying. Each half supports the other: the access half traces the taking, the similarity half compares the use, and together they are how the examination answers the independent development defense. Done together, under the direction of intellectual property counsel, they are built to give the court a chain from the repository event to the competing product, which is what injunctions and damages awards rest on.

What Repository Logs and Developer Workstations Record

Modern source control platforms keep more than most engineering leaders realize. GitHub, GitLab and Bitbucket audit logs record repository clones, fetches and pushes by user and token, the creation and use of personal access tokens and deploy keys, changes to repository visibility and collaborator lists, and the IP addresses behind each action. Those records are where the examiner looks for who pulled the full history of a repository the week before resigning and whether a token survived the offboarding. On the developer workstation, the local .git directory carries its own evidence: the remote origins configured, the reflog of every checkout, the timestamps of the objects and the credential helper entries. Around it sit the usual Windows and macOS artifacts, including shellbags, recent documents, cloud sync client logs and the browser history that can show a personal GitHub account signed in alongside the corporate one.

For engineering data outside of source control, the record is different and just as useful. CAD vaults and product lifecycle management systems keep check out and export logs. Laboratory information systems record query and export events. File servers keep access audit trails when auditing has been enabled, and cloud storage platforms keep download and share events by account. In each case the examiner confirms retention and configuration first, since a platform logs exports when it is set to and keeps audit records for its retention window. We say plainly in the report which sources were available and how each was configured, because a complete account of the record's coverage is what keeps the findings grounded in fact. Robert Knudsen leads this analysis, and the trade secret forensics service page describes the full scope.

© 2026 Global Digital Forensics, Inc. All rights reserved.

Cloud development platforms widen the record further. Continuous integration systems log every build and the credentials that triggered it, artifact repositories record who downloaded compiled packages, and cloud storage holding design files keeps object access logs when they are enabled. Those sources matter because a departing engineer may pull the compiled artifacts or the documentation that explains them in place of cloning the repository. We map all of them in the scoping conversation, rank them by how quickly their retention expires and preserve in that order. The ranking is itself documented, so where a source had already aged out before the engagement the report records the date, and counsel know which questions rest on the remaining sources.

How a Source Code Comparison Runs Under a Protective Order

Code comparison in litigation happens inside rules the parties negotiate, and the engagement is planned around them. The protective order typically restricts the producing party's code to a secure review room or an isolated machine without network access, limits who may view it and governs how notes and printouts are handled. Before anything is compared, we hash the production and the client's own baseline so that the versions examined are fixed and provable. The first analytical pass strips out what both sides would be expected to share: open source libraries, framework boilerplate, generated code and material from published standards. What remains is the candidate region for copying, and it is there that SourceScan and the examiner look for literal matches, structural correspondence after renaming, preserved comments and misspellings, identical unusual constants and, most persuasively, the same bugs reproduced in the same places.

The findings go into an expert report that explains the method before it states the conclusion, because a judge or jury who understands why identical comments and shared defects indicate copying will give the opinion weight. The report ties the comparison back to the access record described earlier, so that the chain reads from the clone event through the workstation to the competitor's product. Depositions in these matters are long and technical, and the source code expert who wrote the report is the one who sits for them. Our work in a cross border intellectual property matter and in a counterfeit product civil seizure supported by imaging and email analysis illustrates how access evidence and comparison fit together, within the factual scope those case studies describe.

The Preparation That Gives a Source Code Case Its Footing

The most valuable preparation is also the simplest: a record of what the company's own code looked like on a given date. Repositories evolve constantly, and a hashed, dated snapshot of the protected version gives the comparison a fixed reference to meet any argument that the matching code was added after the alleged taking. Tagging releases, archiving hashed copies at milestones and keeping the repository audit log exports on a schedule give counsel that reference. The second is individual credentials. A deploy key or service account used by a whole team produces an access record at the team level, so the examiner ties activity to a person through workstation artifacts and badge data. The third is offboarding that retires personal access tokens, OAuth application grants and SSH keys on the day an engineer leaves.

Engineering teams do best when they hand suspected copies to a protocol from the start. A comparison run under counsel's protocol on controlled equipment keeps the record clear of contamination arguments and ties every conclusion to a documented method. Companies with meaningful code value decide in advance who

© 2026 Global Digital Forensics, Inc. All rights reserved.

may touch suspected copies and under what protocol, and keep the source code review and the access analysis separate from the people whose work is in question. Preventive testing matters too. A vulnerability assessment of the repository hosting, the developer VPN and the build pipeline looks for the open doors an outsider would use, and closing them narrows the competing outsider theory when an insider case is argued.

Defend Your Core Intellectual Property and Software Assets

Call Global Digital Forensics at (800) 868-8189 to consult with a senior forensic examiner and source code analyst. The initial consultation is free.

What Repository and Development Logs Record

The guide splits the work into an access half and a similarity half, and the access half depends on what the development platforms recorded. This table lists those sources, what each records and what the examiner confirms about each, so counsel can see which questions the record answers. The third column lists the checks the examiner runs on configuration and attribution. The defense is almost always independent derivation, so the examination traces a path as well as a resemblance.

Repository and development sources, what each records and what the examiner confirms about each.

Source	What it records	What the examiner confirms
GitHub, GitLab or Bitbucket audit log	Clones, fetches and pushes by user, token and address	The platform's logging settings and retention
Personal access tokens and deploy keys	Creation and use of standing credentials	Whether the key is individual or shared by a team
Local .git directory and reflog	Configured remotes and every checkout	The workstation it sits on, tied to its user
CAD vault and PLM system	Check out and export events	That export logging was enabled
Continuous integration system	Builds and the credentials that triggered them	The platform's retention setting
Artifact repository	Who downloaded compiled packages	Downloads as an alternative path to cloning
Workstation artifacts	Cloud sync, recent documents and shellbags	Correlation that ties activity to a person

The garment industry seizure and the counterfeit product matter show onsite imaging and analysis linking communications, inventory and supplier records, within the factual scope each page states, illustrating how access evidence supports an intellectual property claim.

© 2026 Global Digital Forensics, Inc. All rights reserved.

How to Begin and What Happens on the First Call

Our Insider Threat Checklist: Hold the Accounts, Seal the Laptop sets out the items to secure in the order they expire. Bring what you have: the repositories or design systems at issue, who had access to them, when the suspected taking occurred, what the competitor has released and when, and whether intellectual property counsel is already engaged. The examiner asks how source control is hosted and whether audit logging is enabled, how offboarding handles tokens and keys, whether a dated snapshot of the protected version exists and whether any internal comparison has already been attempted. Counsel then receive a written scope covering preservation of the access record, the comparison methodology, the protective order considerations and the timing, with a proposal for each phase. The initial consultation is free, and the hotline at (800) 868-8189 answers 24/7.

Technical References and Further Reading

Earlier reporting on this subject: McAfee and CSIS, Net Losses: Estimating the Global Cost of Cybercrime (archived copy).

Authorities cited in this guide: 18 U.S.C. 1839, Definitions (Defend Trade Secrets Act). Defines improper means and excludes reverse engineering and independent derivation, current codified text at Cornell LII. 18 U.S.C. 1030, Fraud and related activity in connection with computers. Sets the unauthorized access and exceeds authorized access elements of the Computer Fraud and Abuse Act, current codified text. Federal Rule of Evidence 702, Testimony by Expert Witnesses. Governs the admissibility showing for the source code expert's opinions, as amended effective December 1, 2023. NIST SP 800-86, Guide to Integrating Forensic Techniques into Incident Response. Describes the forensic process phases applied to workstation and server evidence, final version 2006.

FREQUENTLY ASKED

Tracing Source Code and Research Asset Theft

Can you show code was copied if the variable names were changed?

Yes, renaming is one of the weakest disguises. Structural comparison looks at control flow, function boundaries, the order of operations and the data structures, none of which change when identifiers are renamed. Comments, string constants, error messages and unusual numeric values often survive untouched. The most telling evidence is shared defects: a bug or an inefficient construction that appears in the same place in both code bases is very hard to explain through independent work. We filter out open source and framework code first so the remaining matches carry weight, and the report explains each category of match in plain terms.

© 2026 Global Digital Forensics, Inc. All rights reserved.

What if the competitor says they developed the software independently?

That is the expected defense, and the examination is built to meet it. Independent development is tested on two fronts. On the access side, the repository audit logs, workstation artifacts and account records are where the examiner looks for a path from the client's code to the competitor. On the similarity side, the competitor's own version history, if produced, shows whether the matching code appeared fully formed on a date that lines up with the access, or evolved through ordinary commits over time. Counsel use those two records together. We present each at its true weight.

Our developers use personal GitHub accounts for work. Does that hurt the case?

It adds steps to attribution and preservation, and the matter proceeds. With personal accounts, the audit log comes from the platform or the individual rather than a company export. Workstation artifacts are where the examiner looks for which accounts were signed in, which remotes were configured and what was cloned and pushed. Counsel can seek the account records through discovery, and the browser and credential helper evidence on the corporate laptop frequently establishes which personal account belongs to which engineer. Moving work to organization owned accounts with audit logging is the preventive fix.

What gets produced in a source code case, and how is it protected?

Protective orders in these matters usually require the producing party to make its code available on a stand alone computer in a secure room, with limits on who may view it, how long, and what notes or printouts may leave. Our examiners work within whatever protocol the court enters. We hash the production on receipt, keep a log of every session and every file opened, and confine notes to what the order permits. Those records matter later, because the producing party will ask how its code was handled and the examiner has to be able to answer precisely.

What determines how long a source code comparison takes?

The size of the two code bases, how many languages and frameworks are involved, how much of the material is open source that has to be identified and filtered, whether version histories are produced or only snapshots, and the review room restrictions in the protective order, which set how many hours a day the code is accessible. A targeted comparison of specific modules moves quickly, and a full comparison of two large products across several releases takes longer. We set the phases and the estimate in writing after the free initial consultation.

Web Addresses for the Sources Cited

1. NIST SP 800-86, Guide to Integrating Forensic Techniques into Incident Response: https://csrc.nist.gov/pubs/sp/800/86/final
2. 18 U.S.C. 1839: https://www.law.cornell.edu/uscode/text/18/1839
3. 18 U.S.C. 1030: https://www.law.cornell.edu/uscode/text/18/1030
4. McAfee and CSIS, Net Losses: Estimating the Global Cost of Cybercrime (archived copy): https://web.archive.org/web/20170102020730/http://www.mcafee.com/us/resources/reports/rp-economic-impact-cybercrime2.pdf
5. Federal Rule of Evidence 702, Testimony by Expert Witnesses: https://www.law.cornell.edu/rules/fre/rule_702

Reviewed by Joseph Caruso, Founder, CEO and CTO, CISSP. Last reviewed November 3, 2026.

© 2026 Global Digital Forensics, Inc. All rights reserved.

Start With a Free Consultation

The initial consultation is free, and requests are handled in confidence. On the first call a senior member of our team listens to the question you need answered, identifies the records and systems involved, and recommends the next step.

Call (800) 868-8189, answered 24/7, request a consultation at evestigate.com/contact, or email info@evestigate.com.

© 2026 Global Digital Forensics, Inc. All rights reserved.