SemanticsAV

An Edge-to-Cloud Al Foundation for Instantaneous and Explainable Threat Intelligence

https://www.semanticsav.ai/

Static Analysis: Syntax Without Semantics

Precision vs Progress

Core Limitation:

Analyzes what code looks like, not what it means

Dynamic Sandbox: The Execution Trap

Observation vs Authenticity

Core Limitation:

Costly execution buys a show the malware directs

Tradition AI/ML:
The Black Box

Power vs Trust

Core Limitation:

Delivers verdicts without verifiable evidence or reasoning

Static Analysis: Syntax Without Semantics

Precision vs Progress

Core Limitation:

Analyzes what code looks like, not what it means

- Fear of false positives restricts detection to the already know
- Reactive by design, a signature is created only after the damage is done
- It analyzes appearance (syntax), completely blind to intent (semantics)
- The slightest obfuscation shatters the model, forcing an endless cycle of updates

Tradition AI/ML: The Black Box

Power vs Trust

Core Limitation:

Delivers verdicts without verifiable evidence or reasoning

Static Analys Syntax Witho

Precision

Core Limita

Analyzes wha like, not what

Dynamic Sandbox: The Execution Trap

Observation vs Authenticity

Core Limitation:

Costly execution buys a show the malware directs

- Organizations accept the cost because execution promises to reveal true behavior
- Minutes per sample creates bottlenecks incompatible with real-time protection
- Full file uploads consume vast bandwidth, imposing strict size limits
- VM detection and time delays trigger benign behavior, bypassing observation completely

ML:

ОХ

r vs Trust

tation:

dicts without

Static Analysis:
Syntax Without Semantics

Precision vs Progress

Core Limitation:

Analyzes what code looks like, not what it means Traditional AI/ML: The Black Box

Power vs Trust

Core Limitation:

Delivers verdicts without verifiable evidence or reasoning

- High detection rates seduce organizations into deploying models they cannot understand
- Opacity forces a binary choice between blind trust and complete rejection
- No path from verdict to evidence leaves every decision unverifiable
- Analysts waste time questioning the system rather than hunting threats

So the real question becomes...

Can we escape the trade-offs?

Imagine threat intelligence with:

- No signatures to maintain
- No hand-crafted rules to tune
- No chasing after threat
- No sandbox theater or waiting
- No blind trust in black-boxes

Attackers shift shape

Defense must understand structure

Zero-Day Detection: Architecture Over Signatures

Coverage AND precision

The Foundation:

Al discovers patterns invisible to human experts

Verifiable Intelligence: Evidence Over Verdicts

Power WITH Transparency

The Foundation:

Geometric position explains the 'why' behind every verdict

Threat Attribution: Lineage Over Labels

Identity AND Ecosystem

The Foundation:

Architectural DNA reveals genetic lineage

Zero-Day Detection: Architecture Over Signatures

Coverage AND precision

The Foundation:

Al discovers patterns invisible to human experts

- ✓ End-to-end learning from raw binaries eliminates all human bias
- Opaque detection model achieves pattern recognition impossible through manual design
- Heavy obfuscation produces distinctive patterns that strengthen detection accuracy
- ✓ Architectural pattern generalization detects threat variants without signature updates

Threat Attribution: Lineage Over Labels

Identity AND Ecosystem

The Foundation:

Architectural DNA reveals genetic lineage

Zero-Day Detec Architecture Ov

Coverage AN

The Foundation

Al discovers pat invisible to huma

Verifiable Intelligence: Evidence Over Verdicts

Power WITH Transparency

The Foundation:

Geometric position explains the 'why' behind every verdict

- Every file positioned against complete code universe
- ✓ Verdict-independent positioning provides parallel validation unconstrained by detection results
- ✓ Geometric neighbors reveal architectural similarities the detection model observed
- ✓ Unfiltered positioning regardless of verdict transforms transparency into confidence metrics

ution:

Labels

D Ecosystem

lation:

DNA reveals

Zero-Day Detection: Architecture Over Signatures

Coverage AND precision

The Foundation:

Al discovers patterns invisible to human experts

Threat Attribution: Lineage Over Labels

Identity AND Ecosystem

The Foundation:

Architectural DNA reveals genetic lineage

- ✓ Identical architecture produces identical DNA regardless of obfuscation or rebranding
- ✓ Genetic positioning maps each sample to specific threat families in code universe
- Evolutionary branching patterns trace how threats mutate and rebrand across campaigns
- Genetic relationships map isolated threats into interconnected campaign ecosystem

Semantics AV Overview





Offline SDK

Zero-Day Protection

Offline Complete

Signature-Free



Cloud Intelligence

Payload-Only Privacy

Threat DNA Mapping

Geometric Proof



Transparent CLI

MIT-licensed

Auditable open source

Exclusive network handler

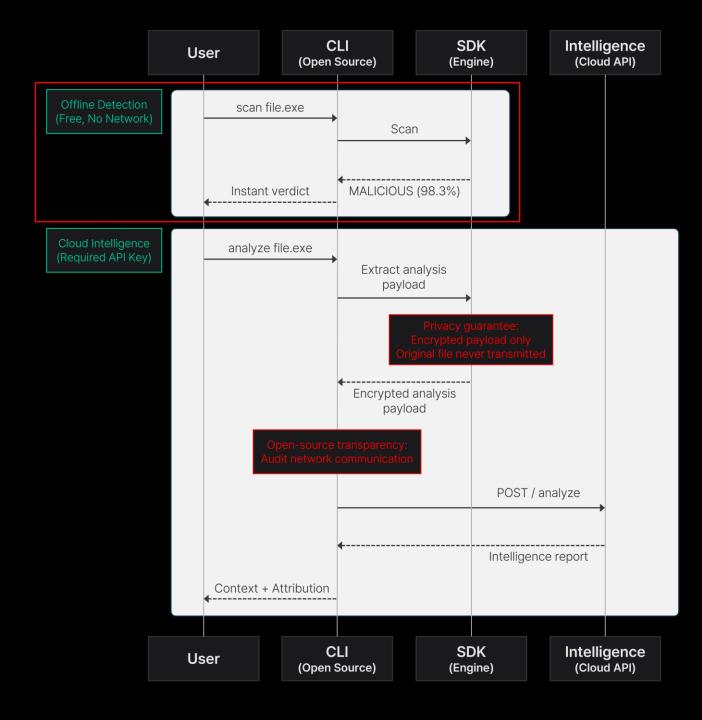
Offline SDK



Al-native detection without signatures or network dependency.

Zero-day protection through architectural pattern recognition; no cloud connection, no compromise.

Complete malware detection in every air-gapped deployment.



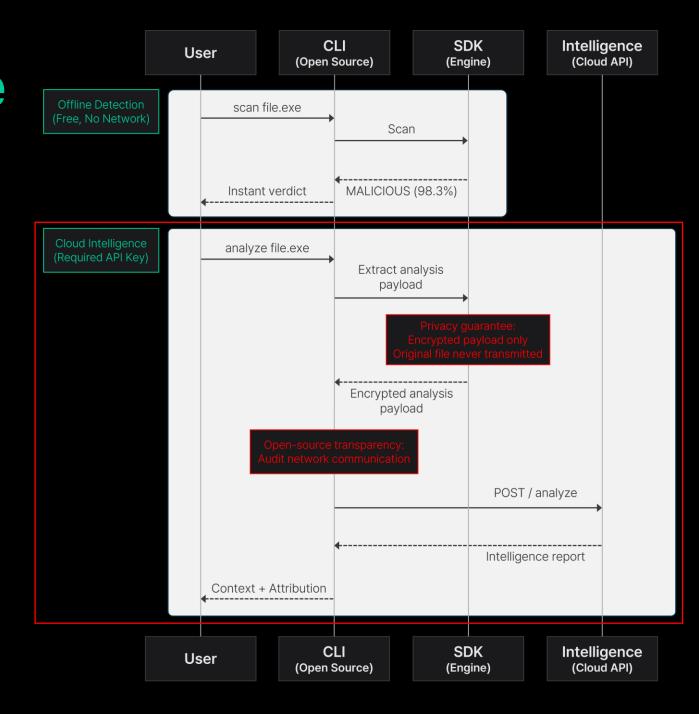
Cloud Intelligence



Threat ecosystem mapping without exposing the file.

Non-reconstructible architectural fingerprinting — positions every sample in the global malware landscape.

Family attribution, campaign correlation, and geometric similarity, instantly.



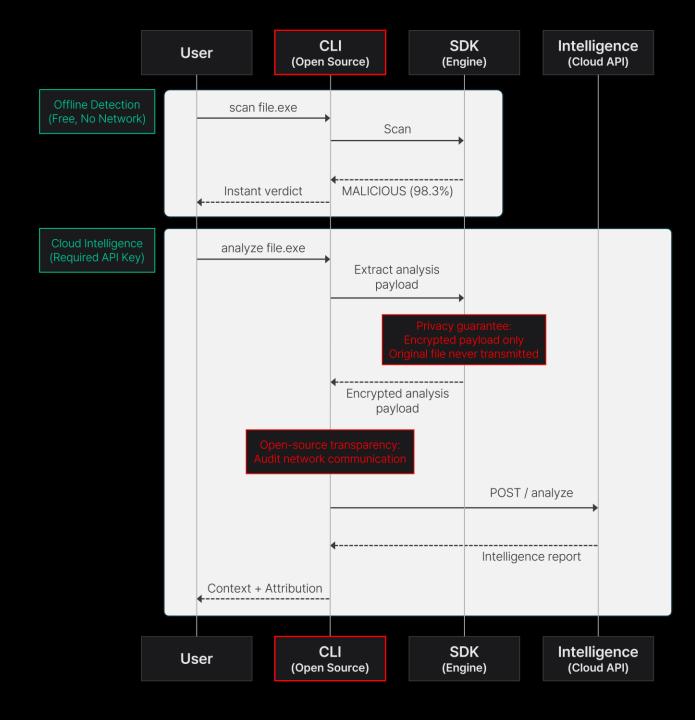
Transparent CLI



Transparency through verifiable architecture.

Every transmitted byte flows through open-source code — auditable, forkable, and independently verifiable

Deterministic payload generation enabling reproducible verification.



Semantics AV Demo

Analyze malware samples instantly with AI-native detection



Enter SHA256 or MalwareBazaar URL

bd7644bfa1c8f6f195ed3d5c2a3ce8b3d69e8cf32d09176ec3ae626c5dcf72b6

Analyze

- Supported: PE (exe, dll), ELF
- Samples must be ≤7 days old

Examples:

bd7644bfa1c8f6f195ed3d5c2a3ce8b3d69e8cf32d09176ec3ae626c5dcf72b6

https://bazaar.abuse.ch/sample/bd7644bfa1c8f6f195ed3d5c2a3ce8b3d69e8cf32d09176ec3ae626c5dcf72b6/

Offline Scanner Performance

Configurations

	SemanticsAV	ClamAV	
Engine Type	Al Model (Signatureless)	Signature-based	
Engine Version	PE: 2025-10-28 05:18 UTC ELF: 2025-10-28 07:46 UTC	Latest version (2025-11-08): 8,724,498 total signatures	

Malware Detection Test

- Dataset: MalwareBazaar daily bulk (2025.10.29 ~ 2025.11.05)
- PE: 701 / ELF: 1,056
- No custom labeling or filtering applied
- Original dataset used as-is for reproducibility
- MalwareBazaar may include benign files

False Positive Test

- Benign Sample Set
- 10,000 files from production systems
- Windows PE: 5,000 (random sampling)
- Linux ELF: 5,000 (random sampling)
- No false positives were detected in either SemanticsAV or ClamAV

Offline Scanner Performance

	SemanticsAV	ClamAV
Total Files Scanned	1,757	1,757
PE Detection Rate	91.58% (642/701)	15.26% (107/701)
ELF Detection Rate	99.24% (1,048/1,056)	75.57% (798/1,056)
Processing Time	37.2s	847.5s (14m 7s)
Throughput	99.5 MB/s	4.5 MB/s
Peak Memory	558 MB	1,782 MB
Average Memory	386 MB	1,496 MB

1 Verdict (Confidence)

SemanticsAV's detection verdict and confidence score

Confidence is identical to the offline scanner

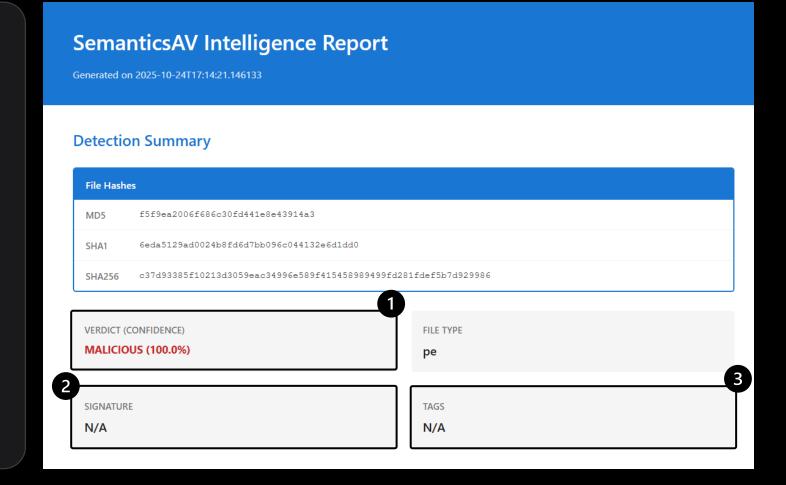
Detection threshold: 0.5

2 3 Signature & Tags

Malware Family Information

Signature and Tag values indicate an exact database match.

N/A indicates no matching sample exists (most common)



(1) Malicious

Samples confirmed as malicious by the platform

2 Suspicious

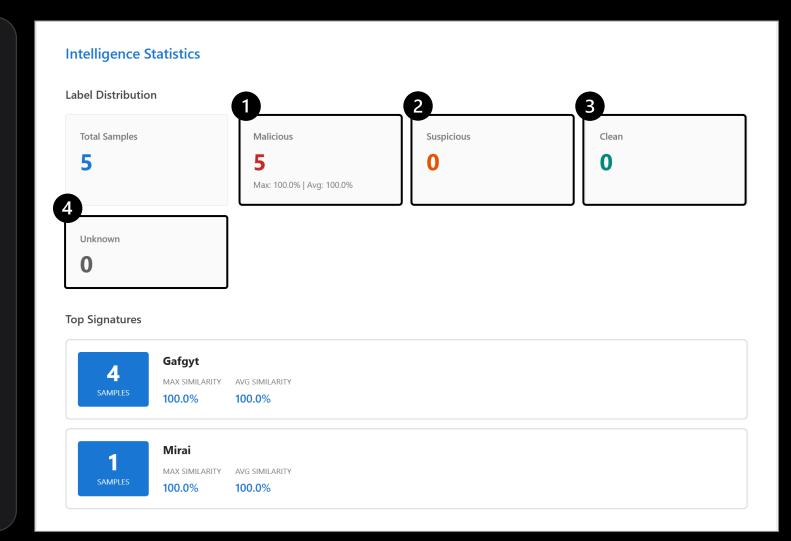
Samples collected from malware feeds but not yet confirmed as malicious

(3) Clean

Samples confirmed as benign by the platform

(4) Unknown

Samples collected from benign sources but not yet confirmed as benign



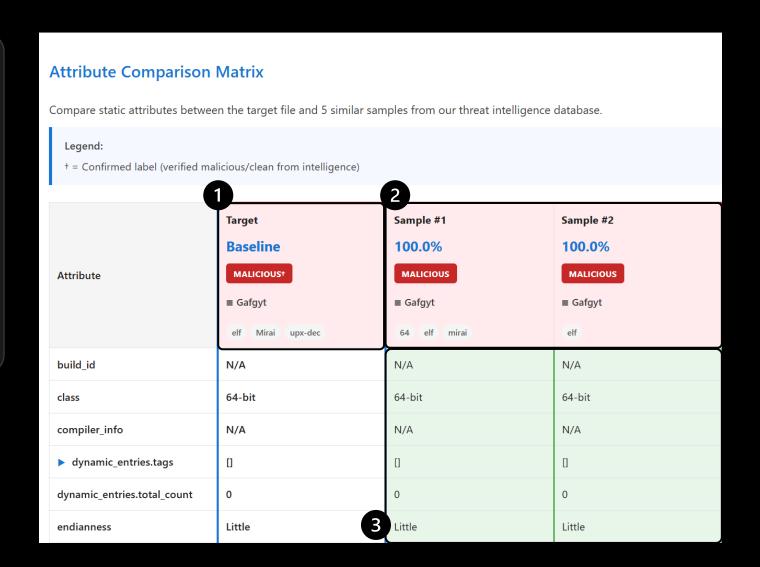
Attribute Comparison Matrix

Compare static attributes between the target file and 5 similar samples from our threat intelligence database.



- 1 Target Information
- 2 Sample Information
- (3) Metadata

Green: Aligned with Target Yellow: Deviation from target



Verdict Summary

LLM-Powered Comprehensive Report Generation

- Technical Analysis
- Threat Intelligence Context
- Recommendations and Response Strategy

COPY REPORT

Executive Verdict Summary

The SemanticsAV Engine has identified this sample as MALICIOUS with 100.0% certainty. Structural analysis reveals a strong genetic link to known Gafgyt and Mirai malware families, indicating its likely use in IoT botnet operations and distributed denial-of-service (DDoS) attacks.

Technical Analysis

The target sample, an ELF executable, exhibits structural characteristics that firmly place it within the Gafgyt family. Analysis of its architectural patterns reveals a design intent focused on network scanning and exploitation, consistent with the typical behavior of this class of malware. Notably, the absence of imported libraries and a high number of exported symbols suggest a self-contained, purpose-built malicious agent designed for efficiency and evasion. Comparative analysis with similar samples in our intelligence database confirms a 100.0% geometric similarity, indicating a direct lineage and likely shared build environment or source code. The presence of the "upx-dec" tag further suggests a common obfuscation technique.

The specific internal structure, including the arrangement of sections and segments, aligns perfectly with known Gafgyt variants. This consistency points to a developer or group utilizing a standardized, potentially modular, codebase for their malicious toolkits. The lack of dynamic entries and the specific entry point further solidify this identification. These structural indicators are not merely superficial; they reflect the core design choices and functionalities embedded within the malware, pointing towards its role in reconnaissance and propagation within targeted networks, likely IoT devices.

#1: Malware Ecosystem Discovery (1)

SemanticsAV discovered structural affinity across multiple malware families

	Target	Sample #1	Sample #2	Sample #3	Sample #4	Sample #5
	Baseline	99.2%	99.2%	99.2%	99.2%	99.2%
Attribute	MALICIOUS†	MALICIOUS	MALICIOUS	MALICIOUS	MALICIOUS	MALICIOUS
	■ a310Logger	■ RemcosRAT	■ a310Logger	■ RemcosRAT	■ RedLineStealer	■ DarkCloud
	exe upx-dec	exe RAT	ехе	exe	exe	exe
compilation_datetime	2025-11-04	2025-10-08	2025-07-22	2025-10-08	2025-07-31	2025-04-24
 data_directories.present 	IMPORT_TABLE, RESOURCE_TABLE, BASE_RELOCATION_T, (4 total)	IMPORT_TABLE, RESOURCE_TABLE, BASE_RELOCATION_T, (6 total)				
debug.by_type.RESERVED	N/A	1	1	1	1	1
debug.entries_count	0	1	1	1	1	1
delay_imports.dll_count	0	0	0	0	0	0

Family	Description
a310Logger	Crypto stealer & keylogger written in Visual Basic
RemcosRAT	Commercial Remote Access Trojan (RAT)
Redline Stealer	Popular Malware-as-a- Service (MaaS) infostealer
DarkCloud	Multi-stage Visual Basic stealer

Key Findings:

- One a310Logger sample mapped to 3 distinct malware families at 99.2% structural similarity.
- Validated via CTI reports (Avast Threat Labs (2021), ANOMALI Threat Research (2021), Rivista Cybersecurity Trends (2022)) showing these families share the same distribution environment and campaign.
- Ecosystem discovery achieved using static-only structural analysis → no signatures, no sandbox required.

#1: Malware Ecosystem Discovery (2)

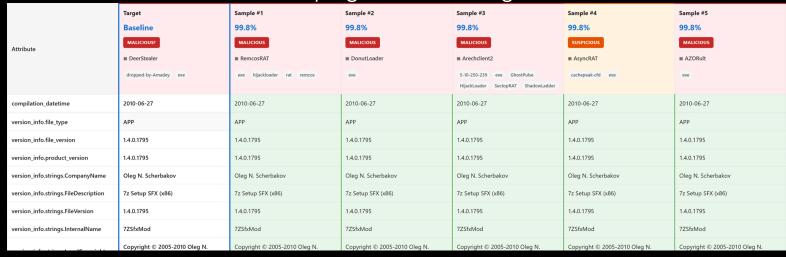
Static Structure Reveals Campaign-Level Linkages

	Target	Sample #1	Sample #2	Sample #3	Sample #4	Sample #5
	Baseline	99.8%	99.8%	99.8%	99.8%	99.8%
Attribute	MALICIOUS†	MALICIOUS	MALICIOUS	MALICIOUS	SUSPICIOUS	MALICIOUS
	■ DeerStealer	■ RemcosRAT	■ DonutLoader	■ Arechclient2	■ AsyncRAT	■ AZORult
	dropped-by-Amadey exe	exe hijackloader rat remcos	exe	5-10-250-239 exe GhostPulse HijackLoader SectopRAT ShadowLadder	cachepeak-cfd exe	exe
compilation_datetime	2010-06-27	2010-06-27	2010-06-27	2010-06-27	2010-06-27	2010-06-27
version_info.file_type	АРР	APP	APP	APP	APP	APP
version_info.file_version	1.4.0.1795	1.4.0.1795	1.4.0.1795	1.4.0.1795	1.4.0.1795	1.4.0.1795
version_info.product_version	1.4.0.1795	1.4.0.1795	1.4.0.1795	1.4.0.1795	1.4.0.1795	1.4.0.1795
version_info.strings.CompanyName	Oleg N. Scherbakov	Oleg N. Scherbakov	Oleg N. Scherbakov	Oleg N. Scherbakov	Oleg N. Scherbakov	Oleg N. Scherbakov
version_info.strings.FileDescription	7z Setup SFX (x86)	7z Setup SFX (x86)	7z Setup SFX (x86)			
version_info.strings.FileVersion	1.4.0.1795	1.4.0.1795	1.4.0.1795	1.4.0.1795	1.4.0.1795	1.4.0.1795
version_info.strings.InternalName	7ZSfxMod	7ZSfxMod	7ZSfxMod	7ZSfxMod	7ZSfxMod	7ZSfxMod
version_info.strings.LegalCopyright	Copyright © 2005-2010 Oleg N. Scherbakov	Copyright © 2005-2010 Oleg N. Scherbakov	Copyright © 2005-2010 Oleg N. Scherbakov			
version_info.strings.OriginalFilename	7ZSfxMod_x86.exe	7ZSfxMod_x86.exe	7ZSfxMod_x86.exe	7ZSfxMod_x86.exe	7ZSfxMod_x86.exe	7ZSfxMod_x86.exe
version_info.strings.PrivateBuild	June 27, 2010	June 27, 2010	June 27, 2010	June 27, 2010	June 27, 2010	June 27, 2010
version_info.strings.ProductName	7-Zip SFX	7-Zip SFX	7-Zip SFX	7-Zip SFX	7-Zip SFX	7-Zip SFX
version_info.strings.ProductVersion	1.4.0.1795	1.4.0.1795	1.4.0.1795	1.4.0.1795	1.4.0.1795	1.4.0.1795

Family	Description
DeerStealer	Subscription-based infostealer sold on dark-web forums
DonutLoader	Open-source in-memory shellcode loader
Arechclient2	Information stealer targeting credentials and system data
AsyncRAT	Open-source modular RAT
AZORult	Credential and cryptocurrency wallet stealer

#1: Malware Ecosystem Discovery (2)

Static Structure Reveals Campaign-Level Linkages



Malware Relationship	Evidence Type	Source
DeerStealer	Delivery	<u>Infosecurity</u>
↔ HijackLoader	Mechanism	Magazine (2025)
RemcosRAT	Campaign	<u>Medium</u>
↔ DonutLoader	Infrastructure	(2025)
HijackLoader	Payload	<u>Red Canary</u>
↔ Arechclient2	Delivery	(2025)
Amadey	Distribution	BlackBerry
↔ AZORult	Infrastructure	Cylance (2020)

Key Findings:

- SemanticsAV uncovered 5 distinct malware families sharing 99.8% architectural similarity.
- Identical 7z SFX build artifacts (2010-06-27 compile timestamp, v1.4.0.1795) across RemcosRAT, AsyncRAT, and multiple loaders indicate a shared MaaS production framework.
- Single-sample static analysis surfaced campaign-level relationships later confirmed by industry reporting (Amadey / HijackLoader distribution infrastructure).
- Rather than being limited to single-sample blocking, analysts can act on the underlying distribution supply chain.

#1: Malware Ecosystem Discovery (3)

SemanticsAV reveals a common-build ecosystem between GhostSocks and SharkStealer

	Target	Sample #1	Sample #2	Sample #3	Sample #4	Sample #5
	Baseline	99.8%	99.8%	99.0%	98.8%	96.2%
Attribute	MALICIOUS* 100.0%	MALICIOUS	MALICIOUS	MALICIOUS	MALICIOUS	SUSPICIOUS
	■ N/A	■ GhostSocks	■ SharkStealer	■ SharkStealer	■ SharkStealer	■ GhostSocks
		exe signed	exe signed	dropped-by-ACRStealer exe signed	dropped-by-ACRStealer exe signed	dropped-by-ACRStealer exe signed
compilation_datetime	2025-07-08	2025-07-08	2025-07-08	2025-07-08	2025-07-08	2025-07-08
file_size	11926976	11961848	11926960	11927072	11926808	11961800
▶ signature.signers	CN=www.gga94ih3nb	CN=www.tjvaq1vztq	CN=www.syv9it0s0b	CN=www.so72a3a478	CN=www.a5scvo6i48	CN=www.1g6z6kxqy3
subsystem	WINDOWS_GUI	WINDOWS_GUI	WINDOWS_GUI	WINDOWS_GUI	WINDOWS_GUI	WINDOWS_GUI
symbols.total_count	0	0	0	0	0	0
tls.callbacks_count	0	0	0	0	0	0
tls.present	Yes	Yes	Yes	Yes	Yes	Yes
version_info.file_type	DLL	DLL	DLL	DLL	DLL	DLL
version_info.file_version	131.110.52845.41465	121.190.52862.58807	57.23.50140.16937	254.10.65534.0	200.50.40000.12340	100.200.30000.0
version_info.product_version	131.110.52845.41465	121.190.52862.58807	57.23.50140.16937	254.10.65534.0	200.50.40000.12340	100.200.30000.0
version_info.strings.						
version_info.strings.CompanyName	Flexera	Flexera	Flexera	Flexera	Flexera	Flexera
version_info.strings.FileDescription	Setup Suite Launcher Unicode					
version_info.strings.FileVersion	131.110.52845.41465	121.190.52862.58807	57.23.50140.16937	254.10.65534	200.50.40000.12340	100.200.30000
version_info.strings.ISInternalDescripti on	Setup Suite Launcher Unicode					
version_info.strings.ISInternalVersion	31.0.24	31.0.24	31.0.24	31.0.24	31.0.24	31.0.24
version_info.strings.Internal Build Number	215864	215864	215864	215864	215864	215864
version_info.strings.InternalName	SetupSuite	SetupSuite	SetupSuite	SetupSuite	SetupSuite	SetupSuite
version_info.strings.LegalCopyright	Copyright (c) 2025 Flexera. All Rights Reserved.					
version_info.strings.OriginalFilename	клипт.ехе	гост.ехе	клипт.ехе	clvJdvkJ2X2QzHLBdzN.exe	клипт.exe	socQ93HWkEQYPxz2fqdZ.exe
version_info.strings.ProductName	ft9bKpoqEX5atVwiAB	43oEYuPBjgb7i9RvbB	dQsyzotwgaCa87A2RG	A2rAoGkUzaKsVYWw8	WMNcndyvozTN2P	KpM3Ure2qoqNQaGsX
version_info.strings.ProductVersion	131.110.52845.41465	121.190.52862.58807	57.23.50140.16937	254.10.65534	200.50.40000.12340	100.200.30000

Family	Description	
GhostSocks	Golang-based infostealer using blockchain-based EtherHiding exfiltrationxy (residential proxy node capability)	
RemcosRAT	Golang-based infostealer using blockchain-based EtherHiding exfiltration	

#1: Malware Ecosystem Discovery (3)

SemanticsAV reveals a common-build ecosystem between GhostSocks and SharkStealer



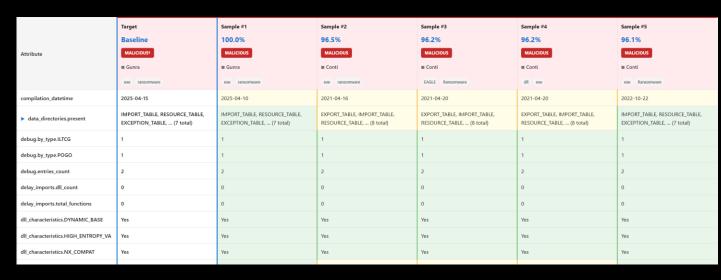
"To the best of our knowledge, no prior relationship between GhostSocks and SharkStealer has been documented in public threat intelligence."

Key Findings:

- Both GhostSocks and SharkStealer samples share identical compilation timestamp, proving simultaneous build from same infrastructure.
- All samples exhibit matching internal build parameters: ISInternalVersion, Build Number, identical CompanyName and FileDescription.
- Structural differences limited to randomized digital certificates and ProductName strings, classic polymorphic evasion technique.
- Implication: Our structural intelligence moves beyond validating known links to defining new malware ecosystems.

2: Evolutionary Lineage Tracking (1)

Automatically revealing lineage, code reuse, and shared ancestry



Conti: A Russia-based ransomware group (2020–2022) whose leaked internal code spawned multiple derivative variants.

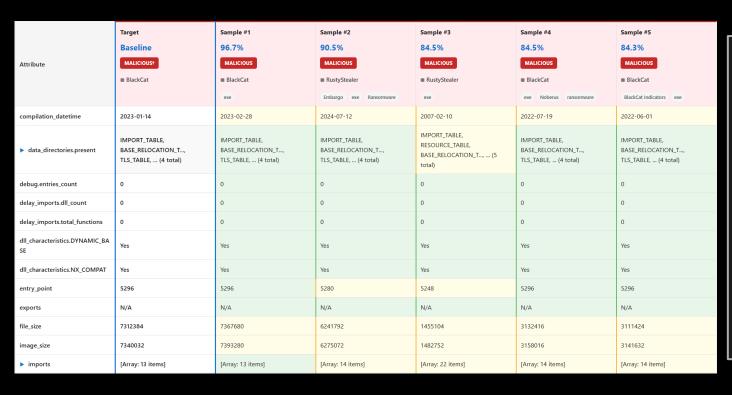
Gunra: A new Apr-2025 ransomware built from leaked Conti's source code using double-extortion (Source: Gunra Ransomware: Conti-Derived Double-Extortion Threat Targeting Global Critical Sectors).

Key Findings:

- A single Gunra query mapped to four Conti variants, automatically validating the leaked codebase relationship.
- The system independently traced the evolutionary lineage from Conti (2021) to Gunra (2025).
- Demonstrates the ability to track malware evolution and code-reuse patterns.
- By exposing evolutionary links between attacks, malware genealogy enables proactive defense that anticipates the adversary's next move.

2: Evolutionary Lineage Tracking (2)

Discovering undocumented ransomware relationships



BlackCat/ALPHV

- Rust-based ransomware-as-a-service (RaaS)
- First observed: Nov 2021
- Use double/triple extortion tactics globally

Embargo Ransomware (Known Relationship)

- Rust-based Raas group, emerged June 2024
- Targets: U.S. healthcare, business services, ...
- Assessed as BlackCat rebrand

(Source: <u>Unmasking Embargo Ransomware: A Deep Dive</u> Into the Group's TTPs and BlackCat Links 2025)

2: Evolutionary Lineage Tracking (2)

Discovering undocumented ransomware relationships



RustyStealer (Undocumented Relationship)

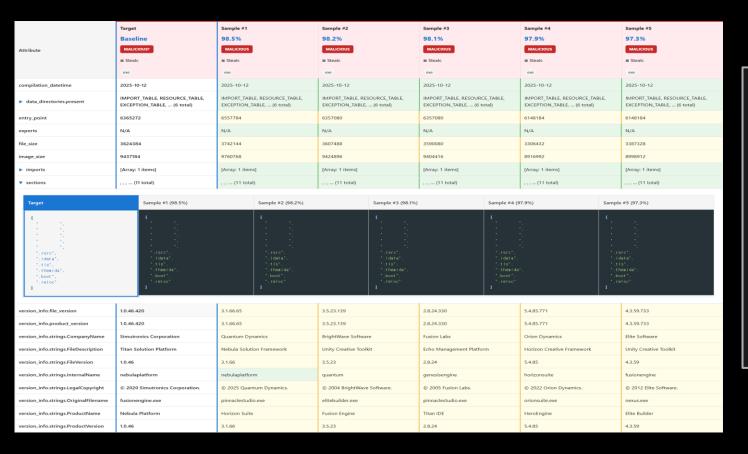
- Rust-compiled information stealer
- Exfiltrates data from browsers, email clients, crypto wallets, password managers
- No publicly documented connection to BalckCat

Key Findings:

- Existing threat intelligence has suggested Embargo-BlackCat relationships through operational similarities.
- Compilation timestamps (BlackCat: Jan 2023 → Embargo: Jul 2024) align with BlackCat's disruption timeline, supporting the strategic rebrand hypothesis.
- SemanticsAV detected 90.5% architectural similarity between the target BlackCat sample and a file tagged as both RustyStealer and Embargo, revealing a previously undocumented connection.
- The triangular relationship (RustyStealer ↔ Embargo ↔ BlackCat) suggests shared development infrastructure
 or coordinated operations rather than coincidental overlap.

#3: Polymorphism as Fingerprint

Stealc Campaign — Evasion Through Polymorphic Identities

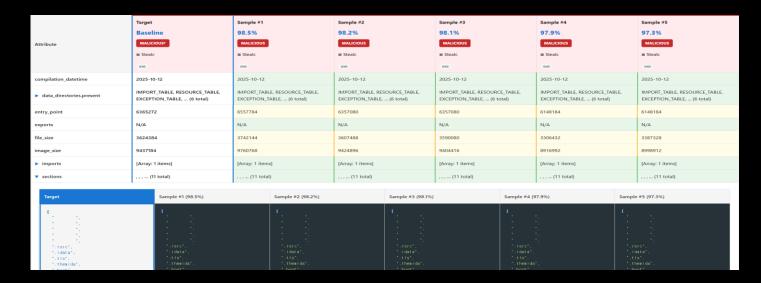


Overview

- Stealc: information-stealing malware targeting browser credentials and crypto wallets
- Themida protection: commercial packer; heavy encryption and anti-analysis make static analysis impractical.
- Polymorphic metadata: per-sample
 CompanyName / ProductName changes to evade reputation-based detection.

#3: Polymorphism as Fingerprint

Stealc Campaign — Evasion Through Polymorphic Identities



"Even when packed with commercial packers such as Themida, SemanticsAV's static analysis can still identify the malware campaign."

Key Findings:

- Attackers easily neutralize traditional defenses by combining commercial packers with trivial metadata tweaks.
- SemanticsAV treats these evasion techniques as persistent structural fingerprints rather than obstacles.
- This undermines the attacker's low-cost repackaging model and forces re-engineering of core malware architecture — increasing attacker cost and complexity.

The Paradigm Inverts

01

Evasion Becomes Evidence.

Obfuscation no longer hides malware SemanticsAV turns evasion into a structural signal

Defenders Automate. Attackers Can't.

02

Defenders generalize from every new sample
Attackers must manually redesign architectures

→ Automation asymmetry becomes defensive leverage

Reuse Dies. Economics Invert.

03

Architectural reuse fails under semantic detection

→ Attackers must pay full development cost for every variant

Q&A