MERKLE SCIENCE

Merkle Science Attribution Framework

Enhancing Explainability and Trust in Blockchain Investigations

2025

Executive Summary

In criminal investigations involving cryptocurrency, attribution is often the first - and most important - step. Who controls a wallet? What type of service does it belong to? Is it linked to known criminal infrastructure or legitimate activity?

This report outlines how Merkle Science equips law enforcement agencies with explainable, auditable attribution data, so that investigators can follow the money with confidence, build defensible cases, and coordinate action across jurisdictions. Attribution is the foundation of blockchain intelligence. Our approach combines multiple evidence sources, clustering heuristics, and validation feedback loops from law enforcement and institutional partners. The result is a system that not only labels wallets but helps uncover the networks and behaviors behind illicit finance.

This case study provides a transparent, end-to-end view of how Merkle Science builds, expands, and verifies attribution on-chain.

- Part 1 outlines the structure and scope of our attribution dataset—how entities are categorized, tags are sourced, and risk levels are assigned using verifiable inputs.
- Part 2 explains how clustering methodologies group addresses into operational structures, revealing how systems behave on-chain and not just who owns what, but how infrastructure operates.
- Part 3 details the validation framework that underpins both tagging and clustering, including our benchmarking process, multi-layered review, and how field use reinforces forensic reliability.

Table Of Contents

1. From Wallets to Entities: Structuring Attribution for Investigations 01

1.1 Entity Classification and Risk Tiering

1.2 Source-Based Tagging and Confidence Tiers

1.3 Mapping Operational Infrastructure

1.4 Tagging Ownership vs. Usage (Cluster vs. Point)

2. Part 2: Attribution Through Clustering 09

2.1 What Is Clustering?

2.2 Clustering on UTXO Chains

- Common Input Ownership (Co-Spend Heuristic)
- Fund Consolidation Heuristic (Multi-In, Single-Out)

2.3 Clustering on EVM-Based Blockchains

- 2.3.1 Deposit Address Clustering
- 2.3.2 Deposit Funder Heuristic
- 2.3.3 Event-Based Clustering
- 2.3.4 Smart Contract Deployer Heuristic

2.4 Behavior Based Clustering

- 2.4.1 Peel Chains
- 2.4.2 Phishing and Drainer Campaigns
- 2.4.3 Transaction Cadence and Timing Patterns

2.5 Custom Heuristics: Entity-Specific Attribution

- 2.5.1 Garantex

2.6 Bridging Off-Chain Intelligence with On-Chain Attribution

3. Part 3: How We Ensure Risk Attribution Accuracy 20

3.1 Prioritizing High-Signal Attribution Inputs

3.2 Structured Validation and Analyst Review

3.3 Metadata Structuring and Oversight Controls

3.4 Attribution Freshness and Ongoing Revalidation

3.5 Feedback Loops and Continuous Model Refinement

4. Conclusion: Attribution That Holds Up 27

MERKLE SCIENCE

Part 1: From Wallets to Entities — Building a Trustworthy Attribution System for Investigations

A process-driven framework for explainable and auditable intelligence

Blockchain attribution creates value only when it directly informs real-world decision-making-whether in criminal investigations, regulatory enforcement, or threat intelligence. At Merkle Science, every attribution entry is grounded in structured evidence, traceable sources, and consistent definitions. We don't just label wallets, we explain who they belong to, how they behave, and why that matters to your case.

Our attribution system is built around three pillars:

- **Structured taxonomy** to standardize how entities are defined and risk is categorized
- **Evidence-based tagging** tied to documented sources and confidence ratings
- **Ongoing validation** through internal review, enforcement benchmarking, and real-world feedback

The sections that follow provide a step-by-step breakdown of how Merkle Science standardizes attribution, starting with how entities are categorized, and and followed by how attribution is applied through evidence-backed tagging and confidence tiering.

1.1: Classifying Entities: What It Is, Why It Matters

In blockchain investigations, attribution is only as valuable as it is interpretable. Freeform or overlapping tags can create investigative blind spots, delay escalation, and introduce risk into enforcement workflows. Knowing a wallet belongs to a custodial exchange is less helpful than knowing what type of custody model it uses or whether it's exposed to sanctioned jurisdictions or mixers. Merkle Science addresses this by applying a standardized, multi-tier taxonomy to every tagged entity.
Entities are classified across three dimensions:

01	**Entity Type**	The high-level function of the actor (e.g., Exchange, Bridge, Mixer, Malware Operator, Wallet Service)
02	**Entity Sub-Type**	Operational specificity (e.g., Custodial vs. Smart Contract Mixer, Spot vs. Futures Exchange)
03	**Risk Tier**	Info, Caution, Medium, High, and Critical based on exposure to illicit activity, legal standing, and transactional behavior [1]

Each tag is mapped to an entity through a layered review of on-chain behavior and off-chain intelligence. We analyze transaction patterns, smart contract interactions, and wallet clustering alongside offchain intelligence such as regulatory filings, licensing disclosures, whitepapers, sanctions lists, domain metadata, and prior enforcement actions.

[1] The risk tiers are shaped by Merkle Science's investigative exposure, including known laundering patterns, actor behaviors, and evolving regulatory enforcement. They are intended as guidelines to assist investigators in prioritizing activity, not as legal or regulatory determinations. We recommend applying these tiers in alignment with your agency's risk appetite, legal mandate, and operational context. Risk levels may be updated as new information becomes available.

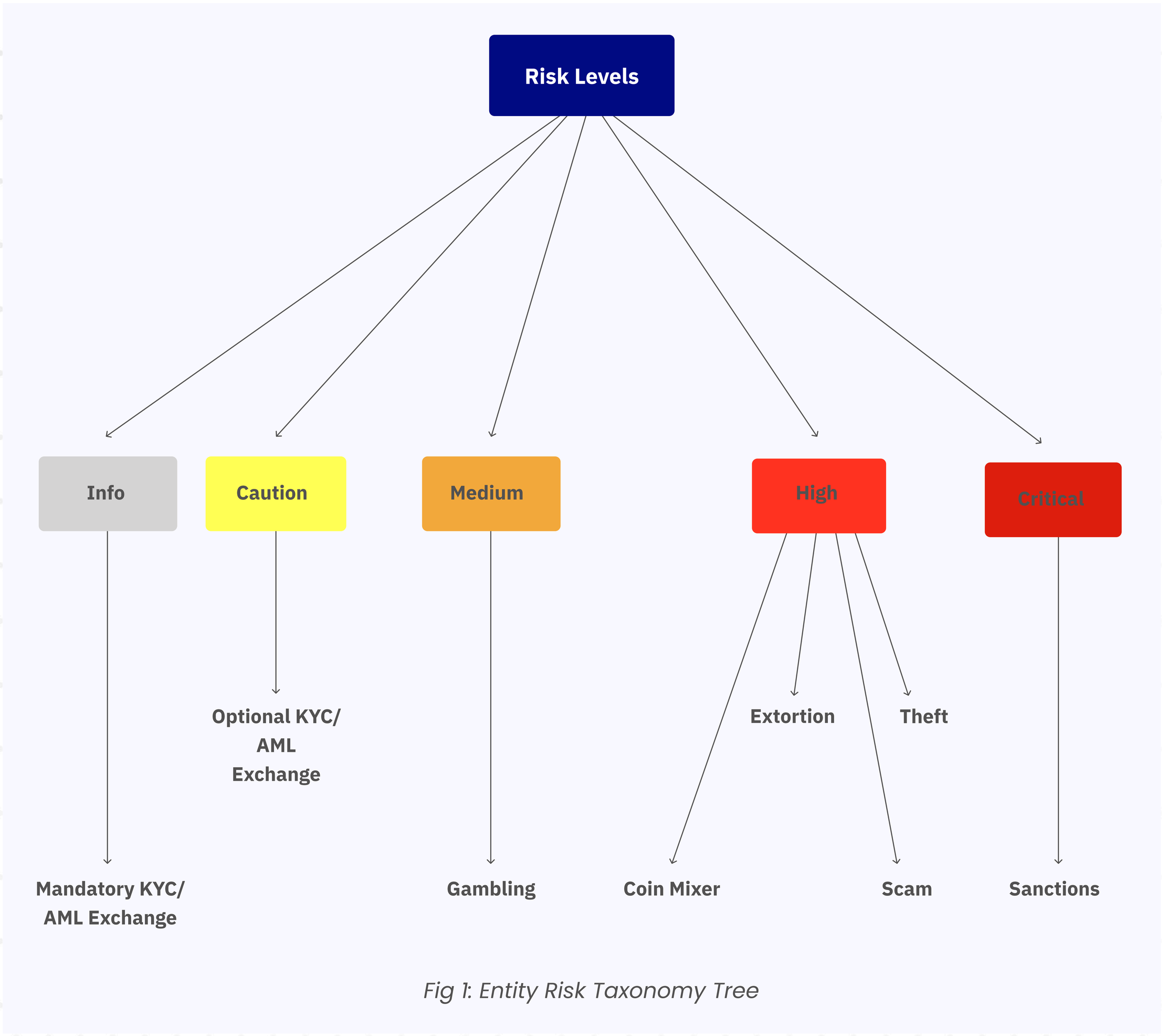


Fig 1: Entity Risk Taxonomy Tree

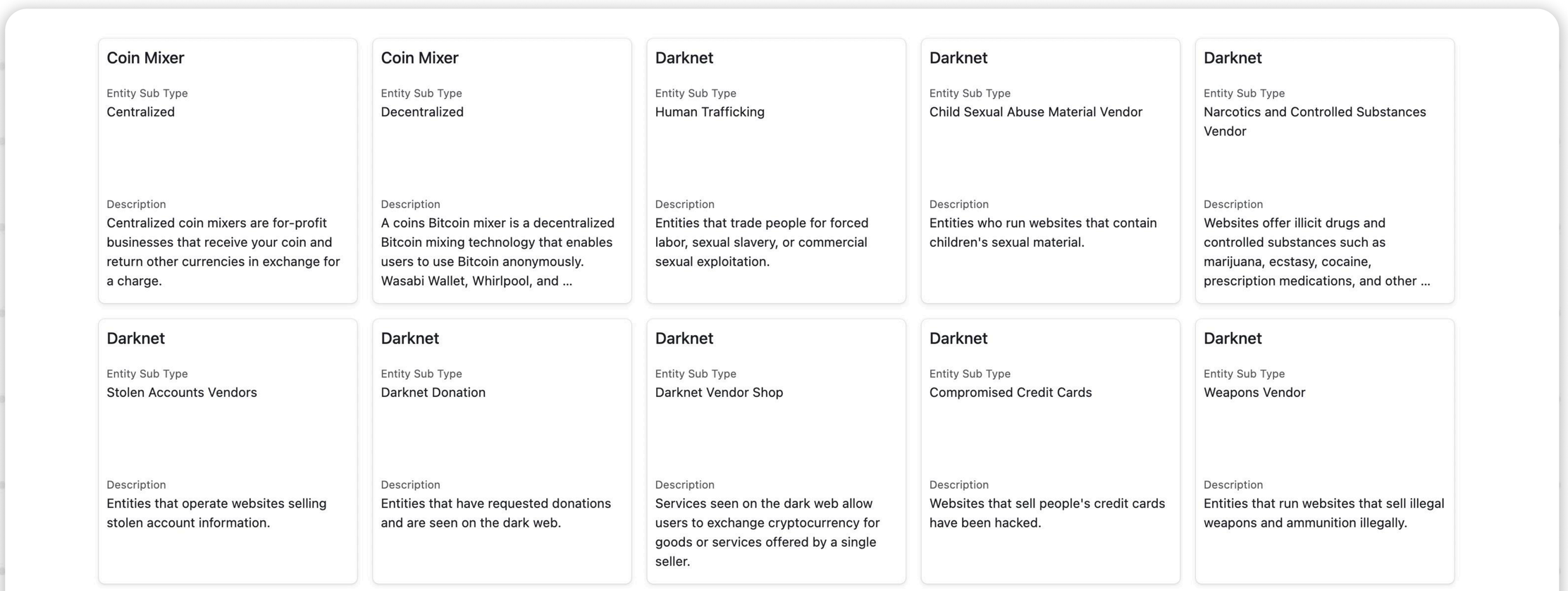


Figure 2: Sample Entity Classifications and Sub-Type Definition

Why this matters?

- Enables structured filtering and triage based on typology, jurisdiction, or case priority
- Supports risk scoring, evidentiary timelines, and enforcement workflows at scale
- Provides investigative clarity across complex or nested infrastructure

1.2: Evidence-Backed Tagging: Attribution You Can Defend in Court

Attribution without evidence is just labeling.

At Merkle Science, every attribution entry is grounded in a documented, reviewable source. Every tag in our system is the output of a controlled, multi-source intake process that prioritizes verifiability, context, and investigative value.

This sourcing model is central to our **evidence-based attribution framework:**

- Each attribution must originate from a **specific credible source type** (e.g., law enforcement disclosure, partner feed, proof of reserves, real time dusting campaigns)
- Each attribution must carry a **confidence rating**, calibrated based on source reliability and cross-validation
- Each attribution must be supported by **verifiable metadata,** including transaction hashes, public disclosures, and timestamped documentation

Note:

→ We do not ingest crowd-labeled tags or accept third-party attribution without independent validation. Instead, we assign attribution only when the data trail is concrete, explainable, and auditable.

Attribution Source Categories and Confidence Tiers [2]

Source	Description	Confidence Tier	Example
Law Enforcement	Seizure lists, indictments, regulatory disclosures	Tier 1 (Confirmed)	DOJ indictment linking BTC wallet to Hydra admin
Partner Feeds	Verified submissions from VASPs, exchanges, or vendors	Tier 1 (Confirmed)	Tagged custodial wallet submitted by compliance team at licensed exchange

Source	Description	Confidence Tier	Example
Proof of Reserves / Posts	Self-disclosed wallets by regulated actors, subject to internal review	Tier 1 (Confirmed)	Binance cold wallet disclosed in proof-of-reserve audit
Dusting Campaigns	Internal tracing tests using micro-transaction	Tier 1 (Confirmed)	Exchange hot wallet identified via traceable dust sent to deposit address
Web Scraping	100+ OSINT sources: Reddit, Telegram, Bitcointalk, GitHub, etc.	Tier 2–3 (Medium–High)	Reddit alias claiming control of scam address, verified via TX hash; or address linked to phishing via public
Customer Submissions	Case-linked annotations submitted via the platform, post analyst review		P2P scam wallet linked to case file
Threat Intel Alliances	Law enforcement, threat intel coalitions, anti-trafficking initiatives	Tier 2 (High)	Blockchain intel from trafficking coalition mapping wallet to CSAM payment

[2] The risk tiers (Tier 1, Tier 2, Tier 3) used in this report are illustrative groupings based on Merkle Science's broader attribution framework. In practice, Merkle Science applies more granular source labels (e.g., confirmed, OSINT, high confidence; suspected) and risk ratings. These tiers are designed to help contextualize confidence levels and attribution sources for investigators. The examples provided are representative, not exhaustive, and may differ from how individual cases are labeled in the live system based on jurisdiction, source validation, or case-specific evidence.

1.3: Infrastructure Mapping: Going Beyond the Deposit Address

Mapping the wallets that power real-world services

Most attribution tools stop at the visible deposit address, without mapping the internal infrastructure. Merkle Science applies tracing methods specifically designed to identify operational wallets, including, hot wallets, treasury addresses, and nested service endpoints used by exchanges, OTC desks, payment processors, and other service providers. These are components critical to understanding how value actually moves within these services.

We focus on this process because **infrastructure attribution is where many attribution models break down.** Without visibility into backend wallet structures, analysts risk false positives, incomplete trails, and misidentified counterparties.

How We do it:

Deposit Dusting

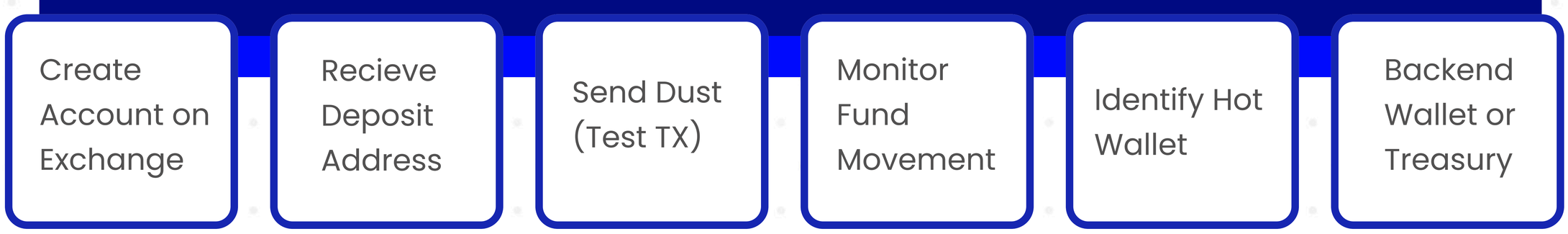
Dusting refers to the practice of sending a negligible, traceable amount of cryptocurrency to a target address to observe how it is handled. In our case, we dust user-facing deposit addresses — unique addresses generated by exchanges or platforms to receive user deposits.

Once the dust is received, it is typically forwarded to the platform's primary hot wallet — a wallet used to manage operational flows like user withdrawals, internal routing, or liquidity provision. By tracking this movement, we can identify the exchange's active infrastructure, including its hot wallets and often-linked treasury or cold storage addresses.

Note:

→ This method is blockchain-agnostic and applies across exchanges, OTC desks, custodians, and other service providers that rely on internal fund routing.

Extended Dusting Flow: From Deposit to Backend Infrastructure



Withdrawal Tracing: When deposit infrastructure is inaccessible, due to login restrictions, inactive accounts, or platform limitations, Merkle Science reverses the tracing process. We initiate a small deposit followed by a withdrawal to a controlled wallet, allowing us to identify the service's active hot wallet or internal routing address.

To ensure the deposit is credited and processed, we use amounts above the platform's minimum threshold. This "operational dusting" allows us to observe how the service handles outbound flows, providing direct attribution of the wallets responsible for withdrawals.

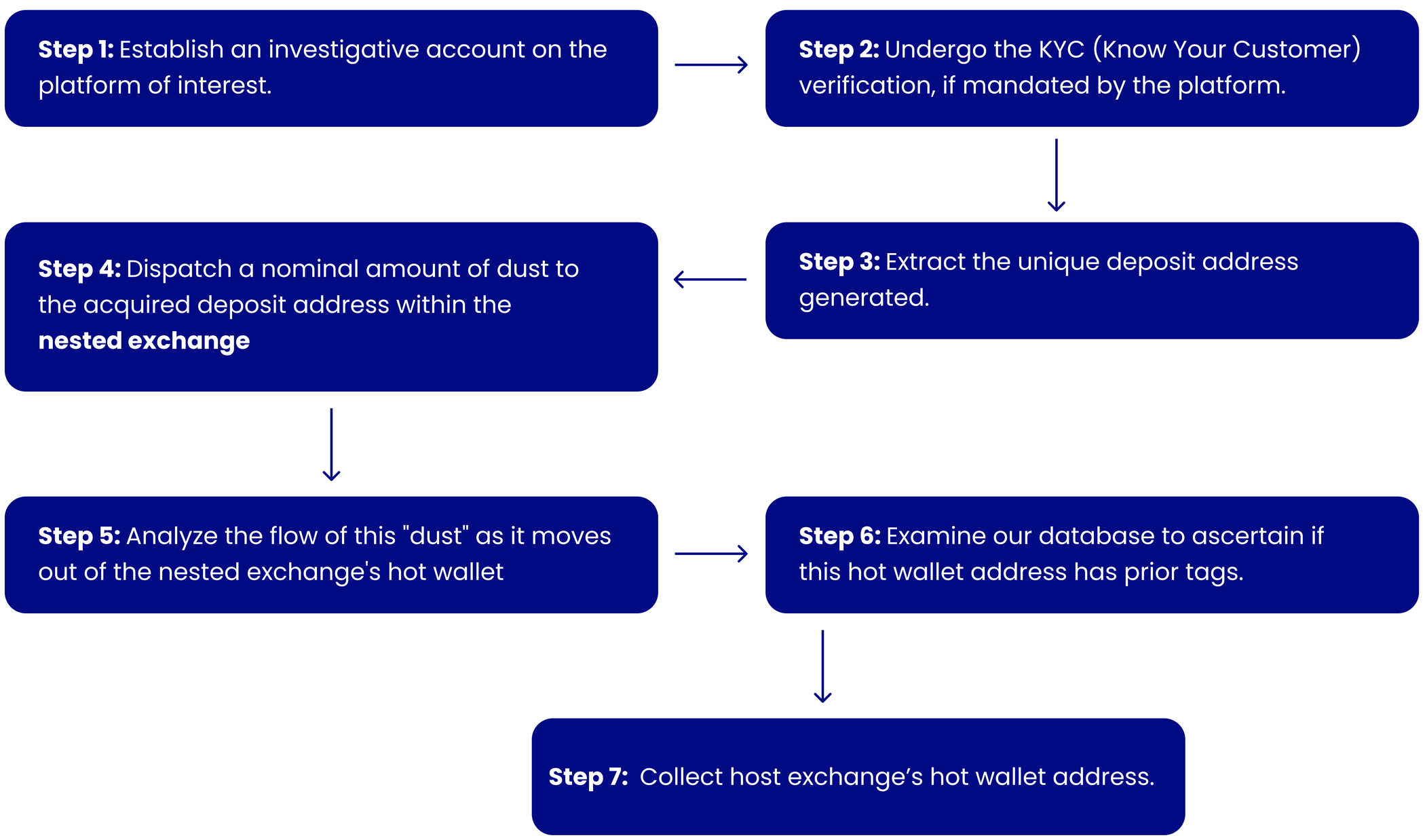
For services that support multiple tokens on the same chain (e.g., USDT and USDC on Ethereum), the process is repeated for each asset to map token-specific infrastructure. Each trace is independently validated before contributing to attribution decisions.

In each case, we aim to perform both a deposit and a withdrawal. This helps us attribute wallet infrastructure more accurately and gives investigators a more complete view of how the platform operates.

Nested Infrastructure Identification: Some platforms operate through accounts on larger exchanges. We detect this by analyzing fund routing patterns, intermediary behavior, and fee structures, allowing us to separate the nested service from its host.

Every identified infrastructure wallet undergoes strict validation process—ensuring that transaction patterns, history, and tagging logic align before it's added to our attribution dataset.

How do we identify a nested service



1.4 Tag Types: Differentiating Infrastructure Control from Usage

Not all addresses on-chain are operated by the entities that use them, attribution must distinguish between who operates an address and who uses it. To reflect this distinction, Merkle Science applies two complementary tagging types: **Owner** (Cluster) Tags and **User** (Point) Tags. These tags clarify whether an address is part of a service's infrastructure or simply used by a third party interacting with it.

Owner / Cluster Tag

Indicates that the address is part of an entity's operational infrastructure—deployed, controlled, or managed directly by the entity. This tag is typically applied to exchanges, issuers, or protocols that maintain persistent access to the address or cluster.

User / Point Tag

Captures the entity interacting with the address, without implying control. This tag is applied when a third party uses infrastructure they don't control, such as a client making deposits to an exchange, a broker operating through a hosted account, or a sanctioned entity transacting through a custodial wallet.

Example: A scammer receiving funds at a Kraken deposit address is labeled the user of that infrastructure; Kraken remains the owner.

This distinction improves investigative accuracy in several ways:

Attribution Clarity

Helps investigators separate control from usage, especially in shared infrastructure environments.

Risk Modeling

Enables more nuanced exposure scoring—highlighting interactions with high-risk entities without misattributing ownership.

Product Labeling

Enables visuals that show layered infrastructure.

Illustrative Scenarios: Owner vs. User Tagging in Practice [3]

Scenario	Owner / Cluster	User / Point	Comment
Sarrafex uses Kucoin for deposits	Kucoin.com	Sarrafex.com	Sarrafex transacts via Kucoin without wallet control; likely hosted or account-level use
Iranian broker hosted on BitForex	BitForex.com	Iranian OTC Broker	Nested use of BitForex with no on-chain control indicators.
OTC desk programmatically accessing Binance wallet	Binance.com	XYZ OTC Desk	Uses internal wallet, but operates as a nested service
White-labeled service on Garantex	Garantex.com	CryptoReal.com	CryptoReal operates through Garantex shared custody
Circle deploys the USDC contract	Circle.com	USDC Contract	Circle deploys and maintains the contract as part of its token issuance infrastructure.
OFAC-sanctioned deposit to Binance	Binance.com	Sanctioned Entity	Binance owns the address, sanctioned party is the point user
Scam wallet receives funds via Kraken deposit	Kraken.com	Scam Wallet	Kraken owns the infrastructure, scammer is the deposit user

[3] **Note:** The examples above are for illustrative purposes only and are intended to demonstrate how Merkle Science applies owner and user tagging in various operational contexts. They do not imply formal investigations, risk judgments, or legal findings against the entities named.

Part 2: Attribution Through Clustering

Clustering connects multiple addresses that are likely controlled by the same entity or infrastructure — enabling investigators to move from isolated wallet tags to a broader operational footprint. It's how we distinguish a single scam address from a coordinated laundering network, or a deposit wallet from a treasury system.

But not all clustering is equal. Over-extended clusters and opaque algorithms can undermine investigations, inflate exposure, and introduce error. Merkle Science applies deterministic, evidence-based clustering designed for investigative clarity and attribution that holds up in court.

2.1 What Is Clustering?

Clustering connects multiple addresses likely controlled by the same actor or service, even when they appear unrelated on the surface. It's what allows investigators to move beyond isolated tags and identify the full operational infrastructure behind an entity or campaign.

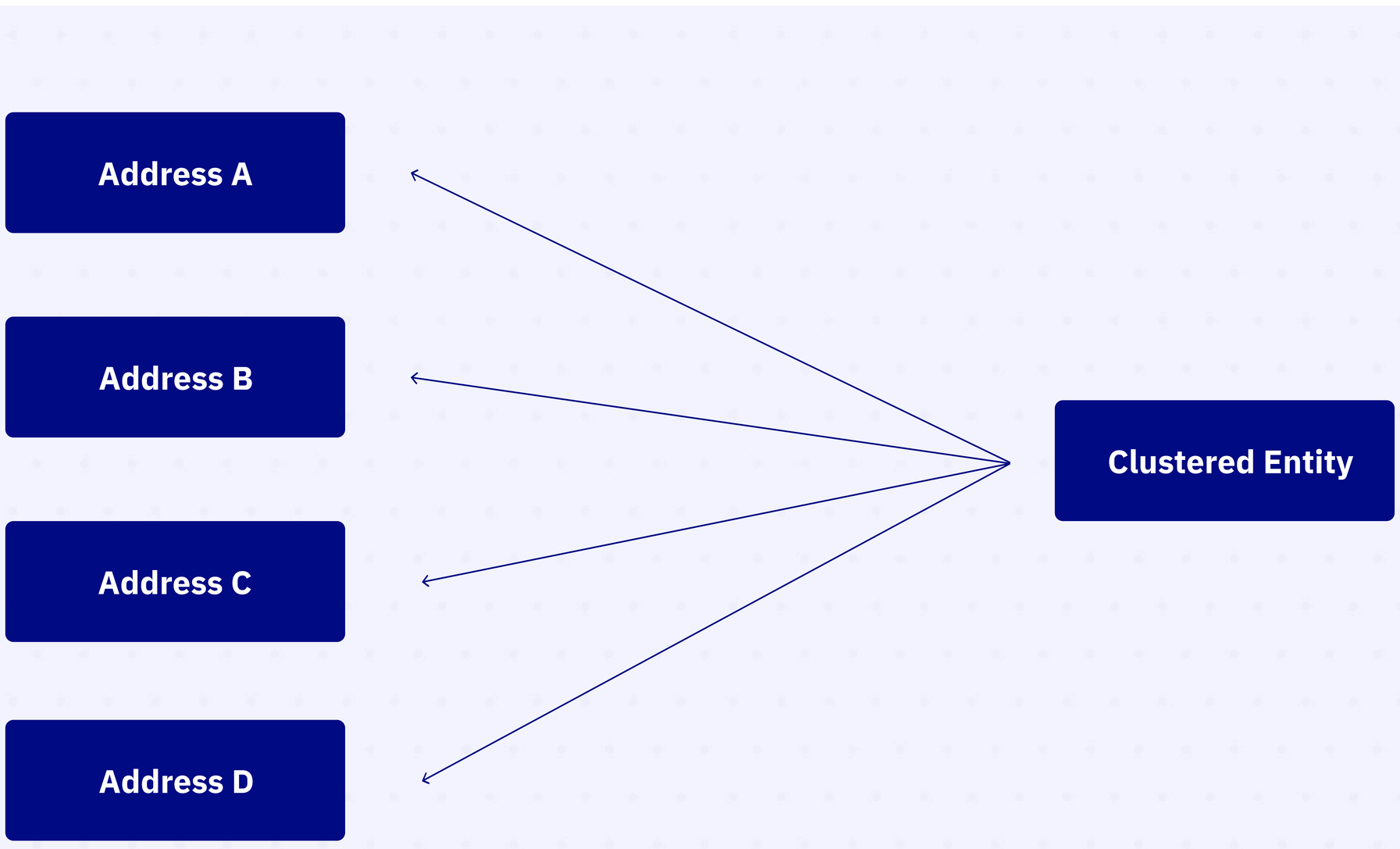


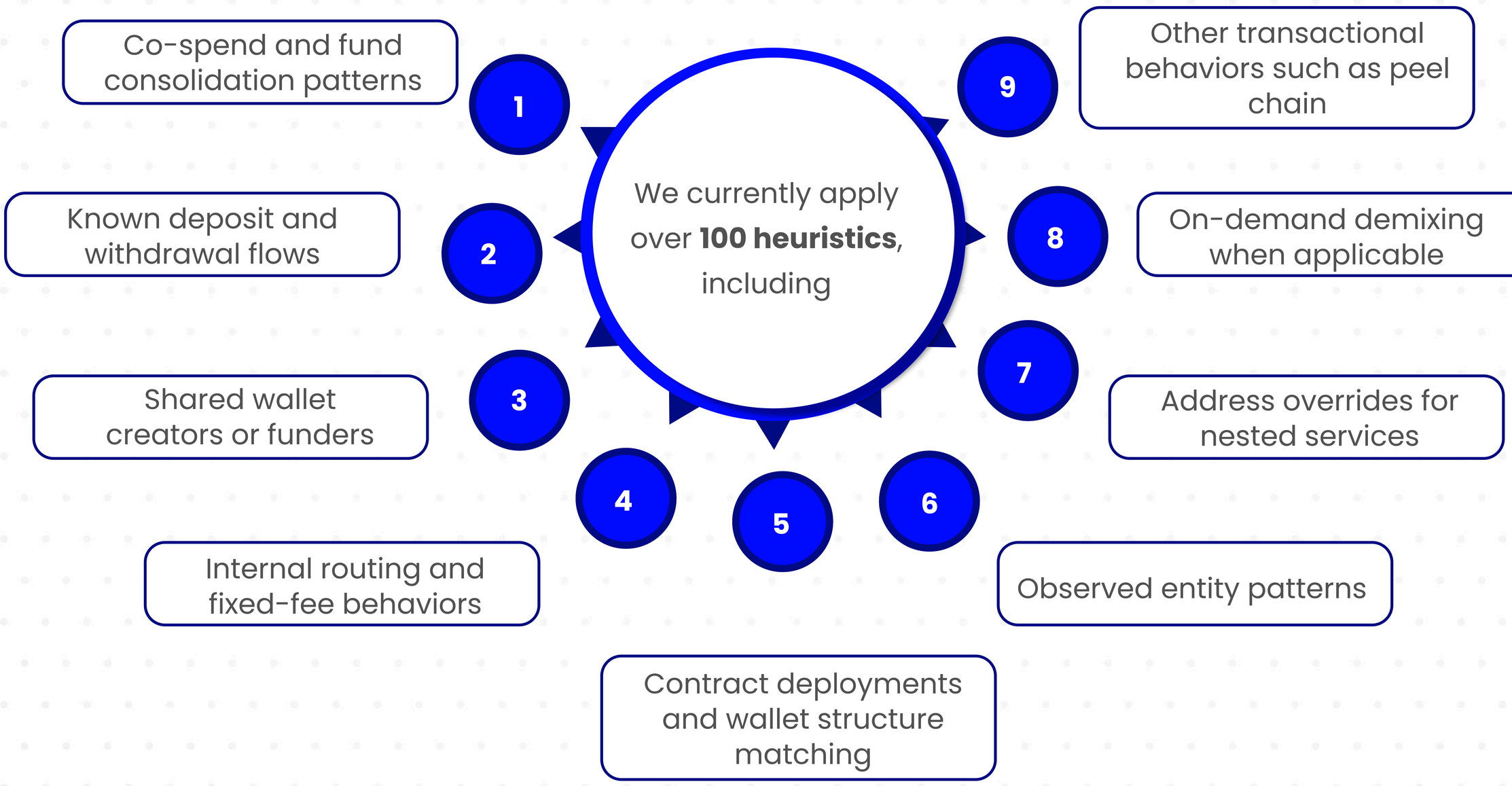
Fig 2: Clustering: Linking Addresses to a Common Entity

Merkle Science anchors clustering in source-backed attribution, not inference. Every cluster begins with a confirmed anchor address—one that is directly linked to a known service through verified sources such as law enforcement disclosures, regulatory actions, sanctions designations, or internal tracing (e.g., test deposits and withdrawals).

From there, we apply a set of rule-based heuristics to identify other addresses that behave in a structurally similar way. A cluster is only extended when the connection is verifiable, consistently reproducible, and supported by on-chain evidence.

We do not use probabilistic clustering or black-box confidence scores. If a link between addresses cannot be clearly explained and defended, it is excluded from attribution.

Our heuristics are tailored to the design of each blockchain, whether it is UTXO-based or account-based. This ensures that both structural patterns and behavioral signals are interpreted correctly in context.The techniques may vary, but the principle is the same: we group addresses only when there's strong, rule-driven evidence that they belong to the same service infrastructure.



When We Cluster—And When We Don't

Every clustering decision must meet internal thresholds for:

- **Repeatability:** Same behavior yields same outcome
- **Traceability:** Every clustered connection must have a documented rationale
- **Auditability:** Analysts and customers can review clustering logic via attribution reports

How We Validate Heuristics

Define → Implement → Code QA → Results QA → Customer Impact Review → Product Migration → Notification

No heuristic is deployed until it passes this full pipeline.

Every rule is tested across multiple chains and use cases, and reviewed for forensic soundness.

We do not expand clusters using heuristic “confidence scores” or unvalidated correlations. If we can’t s how how a connection was made—and why—it doesn’t go into production.

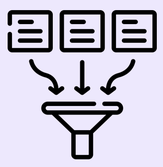
What Makes Merkle's Clustering Unique



Industry-Leading High-Risk Attribution Identify high-risk entity infrastructure up to to 10x faster than other platforms.



Network-Aware Heuristics
Tailored to each chain’s architecture and transaction logic



Ongoing Refinement
Clusters evolve as new behavior or attribution data emerges

Once an address is linked to a cluster, that relationship persists across Merkle Science’s platform uniformly.

2.2 Clustering on UTXO Chains

UTXO chains like Bitcoin and Litecoin represent fund ownership through discrete unspent outputs, not account balances. Each transaction consumes prior outputs and creates new ones, so identifying who controls funds depends on how those outputs are signed and combined.
Merkle Science’s clustering methods are built specifically for this model. Our heuristics account for how services manage UTXOs at scale, detecting patterns like **multi-input spending** and **fund consolidation** to link addresses with shared control.

Common Input Ownership (Co-Spend Heuristic)

On UTXO-based chains, Merkle Science uses co-spend detection to identify when multiple addresses are controlled by the same entity. The logic is simple: if addresses are used together as inputs in a single transaction, each one must be signed by the same private key—implying shared control.

This usually occurs when a wallet doesn’t have one large UTXO to cover a transaction. Instead, it combines smaller outputs from different addresses it controls. We track these patterns across transactions. When a previously unseen address co-spends with a known cluster, we expand that cluster—only if the connection is consistent and explainable.

Example: A 1 BTC payment draws 0.4 BTC from Address A and 0.6 BTC from Address B. Because both were required to complete the transaction, Merkle Science clusters A and B as part of the same operational setup.



Fig:3 UTXO-based Clustering Mechanism

Fund Consolidation Heuristic (Multi-In, Single-Out)

On UTXO-based chains like Bitcoin, services that manage multiple wallets often move funds from several addresses into a single central wallet. This is typically done for internal purposes such as preparing withdrawals, managing treasury balances, or sending assets off-platform.

Merkle Science detects these consolidation patterns by identifying transactions where multiple inputs, associated with the same service, send funds to a single receiving address. When this receiving address appears repeatedly in such transactions, we classify it as a consolidation wallet that is likely part of the service's backend infrastructure.

Example: If deposit wallets A, B, and C all send funds into a transaction, and the only output is Address X, and this pattern occurs across multiple instances, then Address X is tagged as a consolidation point for the service.

> ***Case application:*** *This method can be used in investigations involving* ***sanctioned exchanges and high-risk entities,*** *where front-facing wallets change frequently but internal consolidation remains consistent*

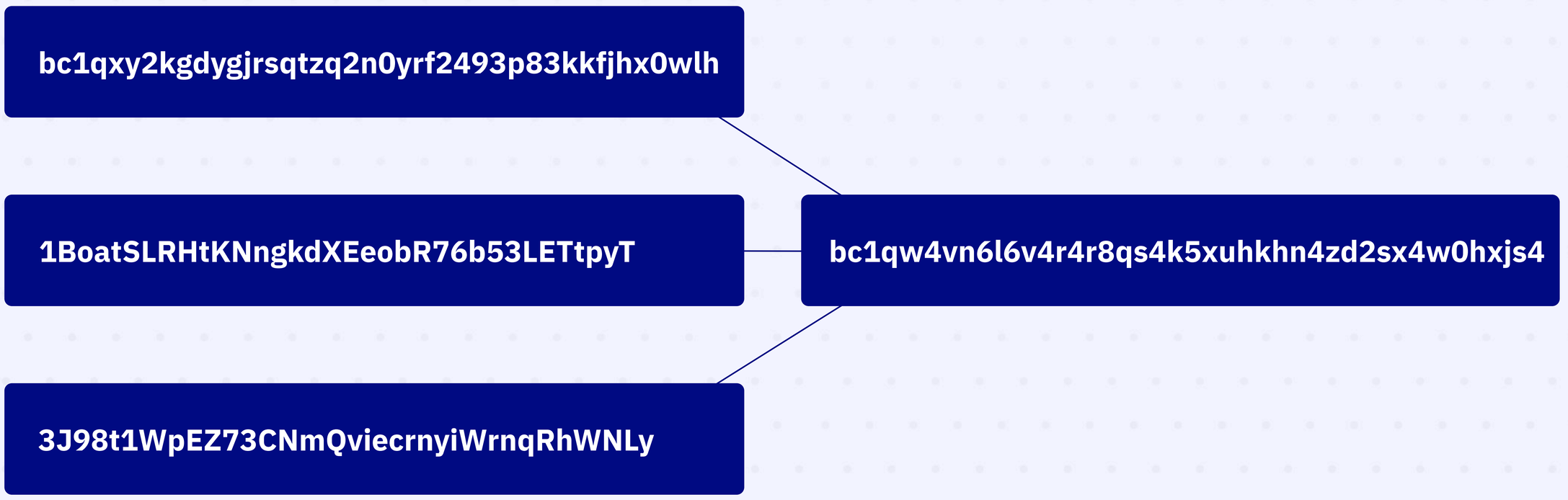


Fig:4 Fund Consolidated CLustering

2.3 Clustering on EVM-Based Blockchains

EVM-based blockchains, such as Ethereum, rely on an **account-based model,** where funds are tracked by user accounts, not discrete transaction outputs like in UTXO-based chains. This model makes it easier to trace ownership at the account level but also requires tailored clustering techniques to account for complex behaviors—like contract interactions, token transfers, and smart contract deployments.

Merkle Science's clustering methods are specifically designed for account-based chains. We take into consideration the **account architecture**, **transaction patterns, and contract interactions** that are unique to these ecosystems. This ensures we can accurately map addresses and detect entity infrastructure in ways that account for the dynamic and often fragmented nature of account-based blockchains.

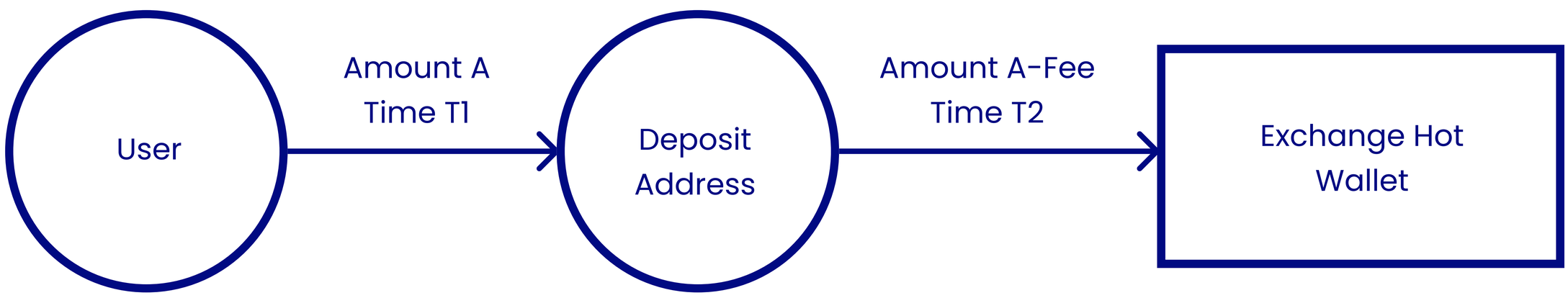
2.3.1 Deposit Address Clustering

On EVM-based platforms, services often generate unique deposit addresses for each user. However, some entities reuse the same deposit address across multiple transactions.
At Merkle Science, we monitor how funds move from these deposit addresses. When we observe that multiple deposits consistently flow into the same destination wallet, we classify that wallet as the service's primary hot wallet.

Once the hot wallet is confirmed, we apply clustering heuristics to identify other deposit addresses connected to it—effectively mapping the intake infrastructure of the platform.



This heuristic helps attribute front-end activity to core operational wallets, even in cases where deposit address reuse is limited.



2.3.2 Deposit Funder Heuristic

On account-based chains like Tron, many centralized exchanges use deposit addresses that only receive non-native tokens, such as USDT (ERC-20). Because these tokens require native gas (ETH) to be transferred, exchanges maintain a group of internal wallets—**deposit funders**—responsible for supplying ETH to these deposit addresses so they can forward the tokens to the platform's hot wallet.

At Merkle Science, we identify these deposit funder wallets through a combination of dusting techniques, graph analytics, and long-range pattern detection. Once a deposit funder is confirmed, we monitor its outgoing transactions to identify addresses it consistently tops up.

When the same funder repeatedly sends ETH to new addresses that immediately execute ERC-20 transfers, we apply additional clustering logic—including fund flow analysis, transaction timing, and behavioral pattern matching—to confirm that these addresses are indeed deposit addresses operated by the same entity.

Case Application:

Merkle Science has used this heuristic to map wallet and service provider infrastructure, including identifying deposit networks behind exchanges and custodial services operating on behalf of users. In many cases, the deposit addresses remained hidden from traditional token-flow methods but were visible through consistent gas provisioning patterns.

2.3.3 Event-Based Clustering

Event-based clustering links addresses that consistently appear together within the same **on-chain event or operational context.** Unlike heuristics focused solely on fund movement or wallet ownership, this method captures addresses that act in coordination during structured events—such as contract executions, token launches, exploit flows, or gasless transaction batches.

At Merkle Science, we apply event-based clustering when:

- Multiple addresses interact with the same contract in a single transaction or batch
- Wallets execute coordinated actions within the same block or defined time window
- Repeated behavioral patterns emerge—such as flash loans, token approvals, or liquidity pulls
- Exploit or scam wallets operate using the same infrastructure or playbooks

By anchoring attribution to shared event structures—not just transfer logic— we can identify **ephemeral or burner wallets** that would otherwise remain siloed. Ephemeral or burner wallets are temporary addresses created for short-term use often to execute a single transaction, obfuscate ownership, or avoid detection before being abandoned.This method is especially useful for surfacing **campaign-level coordination**, **infrastructure reuse,** and **DeFi actor attribution**.

Use case: *A simple yet effective way to attribute infrastructure linked to DEX activity or identify wallets operating within coordinated exploit campaigns.*

2.3.4 Smart Contract Deployer Heuristic

On EVM-based chains, every smart contract is deployed by a wallet address—called the deployer. This heuristic links contracts to the entity that controls the deployer, allowing Merkle Science to attribute infrastructure even when names or interfaces vary.

We apply the deployer heuristic in two stages:

Primary Attribution	We tag a deployer address with an entity label based on verified sources—such as law enforcement disclosures, official posts, or internal validation (e.g., a DEX confirming its factory deployer).
Heuristic Application: Cross-Chain Propagation	Once a deployer is tagged on one chain (e.g., Ethereum), we apply a rule: all contracts deployed by that same address on EVM-compatible chains (e.g., BSC, Polygon, Avalanche) inherit the same attribution. This propagation assumes control continuity across chains unless contradicted by behavioral differences.

Use case: *Enables investigators to map a DEX's complete contract ecosystem, spot cross-chain clones, or link scams deploying from the same wallet.*

2.4 Behavior Based Clustering

Behavior-based clustering links addresses that exhibit consistent on-chain behaviors patterns that may not be visible through ownership or flow alone, but emerge through how wallets interact, transact, and evolve over time.

At Merkle Science, this form of clustering is especially valuable when:



Note:

→ Custom and behavioral heuristics go through the same validation pipeline as standard rules. They're deterministic, rigorously tested, and produce the same output quality—consistently reproducible clusters based on field-proven patterns.

We analyze **execution patterns, timing intervals, transaction composition**, and more to uncover operational links between wallets that are otherwise disconnected at the surface level.

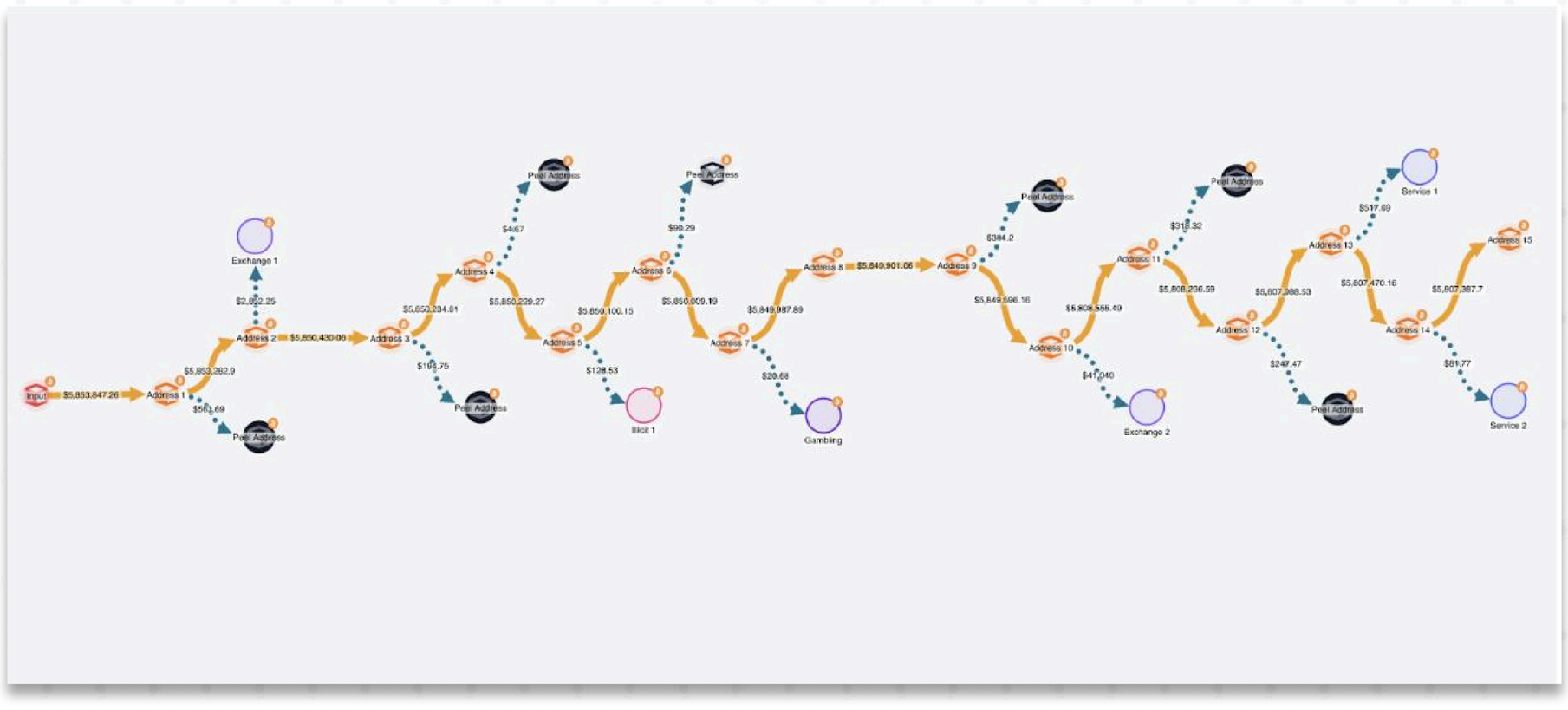
2.4.1 Peel Chains

In a peel chain, funds are moved through a sequence of addresses, each retaining a small remainder. This tactic is commonly used by laundering operations to fragment value and complicate tracing.

Merkle Science detects this structure by tracking:

- Progressive value reduction across addresses
- Consistent spacing in block intervals
- Reused scripts or interaction methods

Once identified, we cluster the intermediate nodes as part of a laundering system—often linking them back to a central entity or consolidation point. Peel chains are especially well-suited to **graph-based clustering,** where each transaction forms a directional edge and reveals a linear flow pattern. By combining value reduction behavior, repeating timing intervals, and structural graph paths, Merkle Science can trace and attribute even long, obfuscated peel chains to their origin or end-point entity.



Visual Depiction of Chain Peeling (Source: Tracker)

2.4.2 Phishing and Drainer Campaigns

Wallets used in phishing campaigns often interact with the same dApps or contracts, deploy front-running scripts, or execute token approvals in repeatable sequences. Even if hosted across new addresses or chains, these behavioral fingerprints remain.

Examples of repeated behavioral patterns that we analyze:

1.
Contract interaction sequences

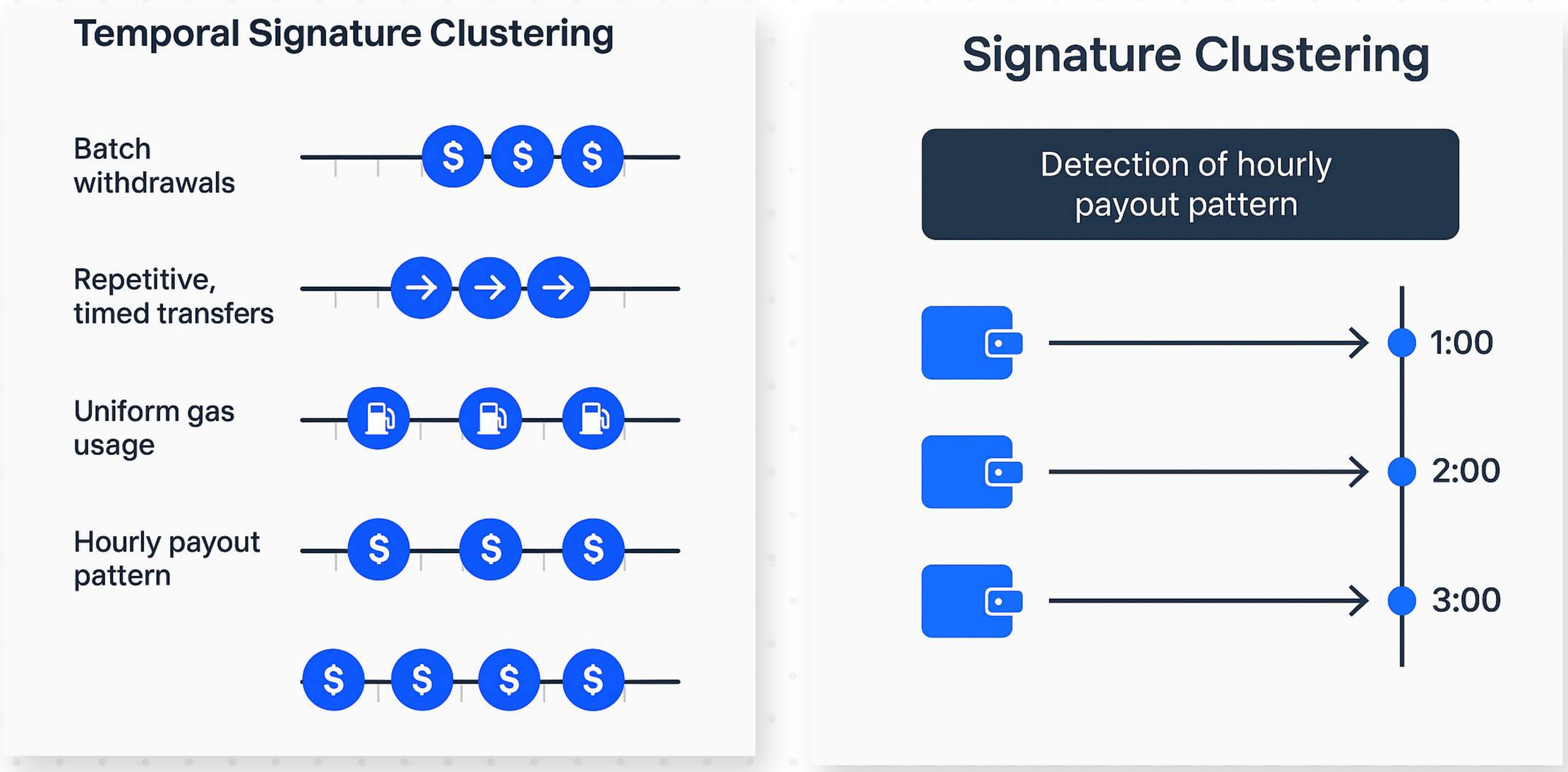
2.
Token approvals to malicious contracts

3.
Gas funders reused across campaigns

This allows us to surface new scam infrastructure early—even if the address is freshly created or cross-chain.

2.4.3 Transaction Cadence and Timing Patterns

Behavioral clustering also accounts for when transactions occur—not just how. We detect temporal signatures like batch withdrawals within tight block intervals, repeating hourly transfers, or uniform gas usage across wallets. These timing patterns serve as behavioral fingerprints, enabling us to cluster wallets used in laundering pipelines, arbitrage systems, or automated payout structures.



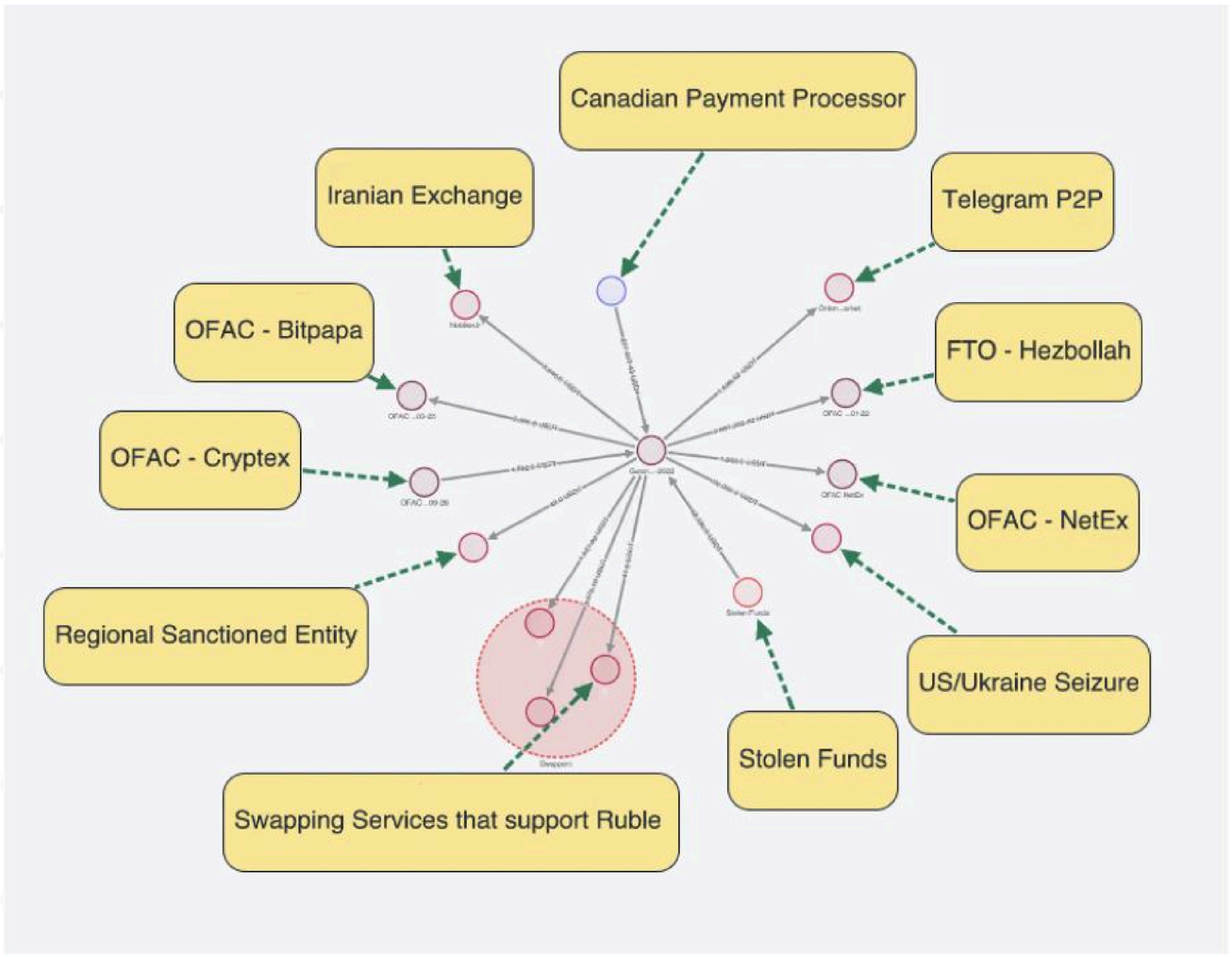
Wallets initiating synchronized hourly transactions suggest automation or shared infrastructure—key indicators used in temporal signature clustering.

2.5 Custom Heuristics: Entity-Specific Attribution

In addition to universal clustering logic, Merkle Science develops custom heuristics for high-risk or structurally unique entities—where standardized methods fall short.

2.5.1 Garantex

Merkle Science maintains a proprietary behavioral model tailored to Garantex, based on field investigations and long-range transaction analysis. Key tactics include:



Cluster View: Garantex Infrastructure and Links to Illicit Entities

High-Frequency Hot Wallet Rotation:

Garantex routinely rotated its hot wallets (approximately every five hours), complicating tracing efforts significantly. The rapid turnover disrupted address clustering analysis, allowing funds to move without being flagged by conventional risk-scoring models.

Layered Wallet Handovers:

Funds were deliberately transferred through multiple intermediary wallets in intricate, multi-layered transaction paths, intentionally fragmenting transaction histories to hinder traceability.

Fixed Gas Fee Deposits:

Depositors consistently prefunded fixed amounts designated specifically for gas fees, a methodical laundering hallmark. This regular and predictable behavior cleverly evaded automated transaction monitoring systems.

With Merkle Science's industry-leading coverage of sanctioned actors on Tron and our extensive High Risk Entities (HRE) dataset—comprising millions of addresses, including 250+ unique entities not covered elsewhere—investigators can uncover, link, and escalate illicit activity with precision, enabling faster case development and proactive disruption of laundering operations.

2.6 Bridging Off-Chain Intelligence with On-Chain Attribution

Merkle Science's attribution system does not rely solely on blockchain data. It combines on-chain behavioral patterns with verified off-chain intelligence to provide a fuller, more legally and operationally meaningful picture of risk. It incorporates metadata fields that offer deeper insight into how an entity operates, where it is based, and how it relates to evolving regulatory obligations. These fields support more targeted investigations, exposure mapping, and compliance analysis.

Key metadata fields include:

Location: Identifies the country where the entity is primarily incorporated, based, or operates. This helps assess jurisdictional risk and legal exposure across global investigations.

Custody Level: Classifies entities by how they manage user funds—Custodial, Non-Custodial, Hybrid, or Decentralized. This informs risk scoring, licensing obligations, and AML traceability.

Entity State (is_active): Reflects whether an entity remains operational. This is useful when distinguishing between active services and dormant or seized infrastructure that may still receive funds.

Owner vs. User Attribution: At the address level, Merkle distinguishes between the entity that owns infrastructure (e.g., Circle) and the service that uses it (e.g., platforms issuing USDC). This helps uncover nested services and shared custodial flows that may otherwise remain hidden.

Effective Date (attribution_effective_date): Captures the point in time when an entity was formally flagged for risk—such as being sanctioned, seized, or linked to illicit activity. This timestamp supports retrospective exposure analysis, audit trails, and historical backtesting across investigations and compliance reviews.

High-Risk Fiat Support: Merkle Science flags also flags entities that support fiat currencies tied to elevated sanctions or geopolitical risk. This includes, but is not limited to, currencies such as:

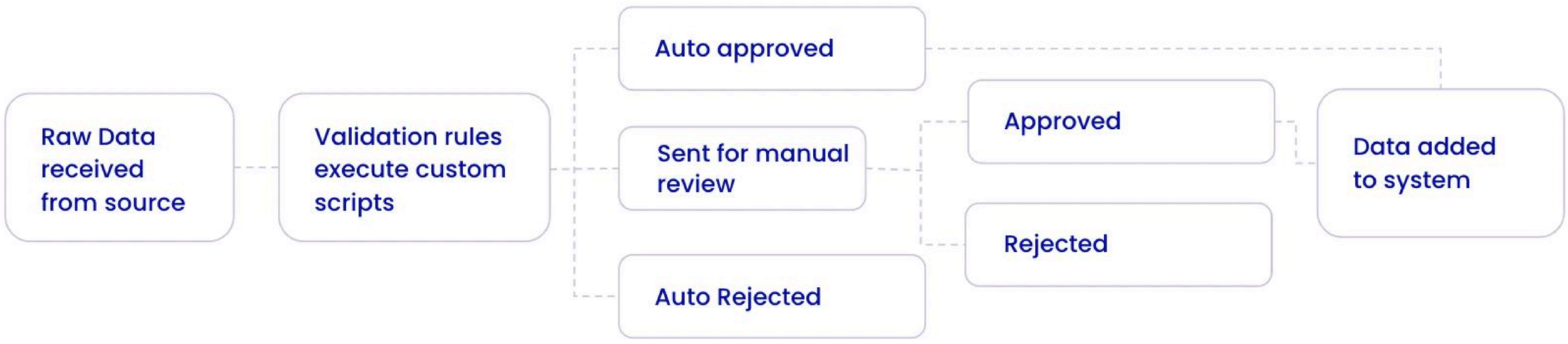
- Russian ruble (RUB)
- Iranian rial (IRR)
- Venezuelan bolívar (VES).

Part 3: How We Ensure Risk Attribution Accuracy

In blockchain investigations, attribution is only as valuable as it is accurate. Misclassification doesn't just slow down a case—it can compromise an entire line of inquiry, mislead enforcement partners, or erode trust in the data. That's why, at Merkle Science, we treat attribution not as a static label, but as a living, auditable claim.

This section breaks down how we ensure our attributions are trustworthy from start to finish and how we collect data, validate it, assign risk, and monitor it continuously as the underlying behavior evolves. We don't just aim to be accurate once—we aim to stay accurate as the chain moves forward.

What follows is a clear, step-by-step explanation of the system behind our intelligence. From source selection to analyst review to real-time updates, you'll see how Merkle Science ensures that every cluster, tag, and risk score stands up to investigative scrutiny.



Data Quality Assurance Snapshot

3.1 Start with High-Signal, Actionable Data

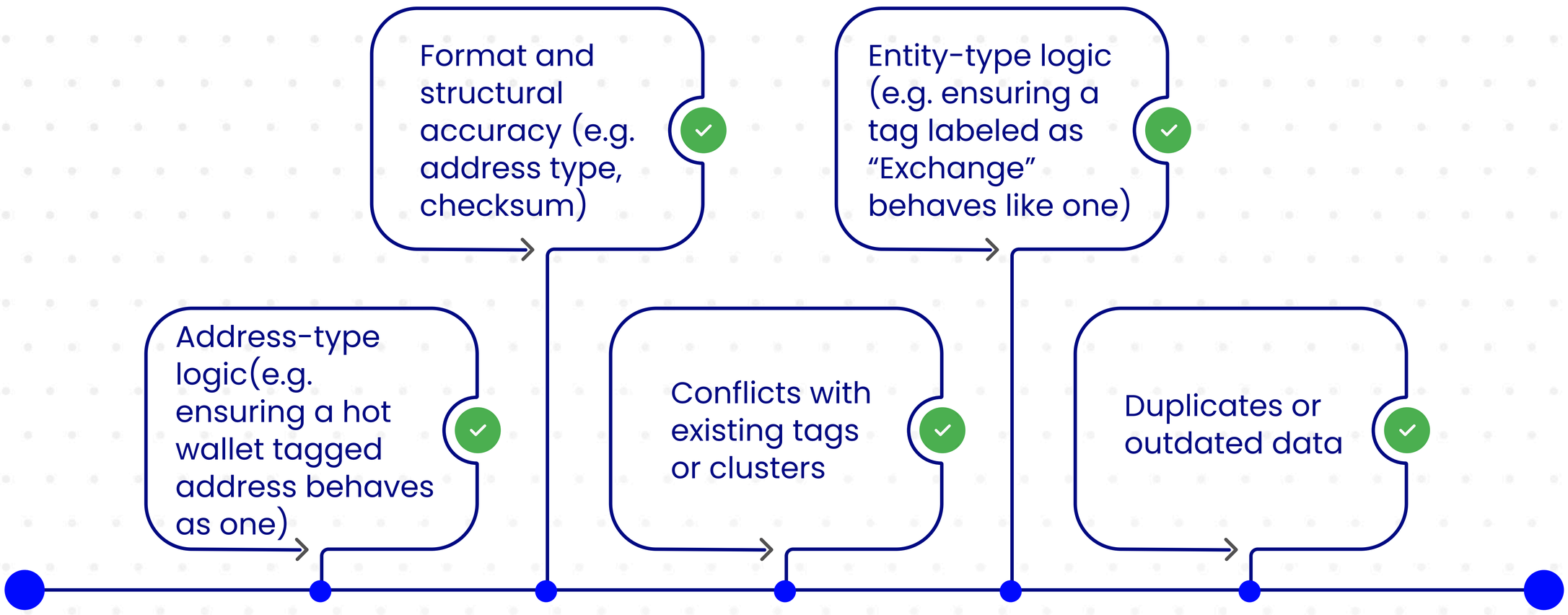
We've already covered the importance of reliable sources and how we collect data in Part 1. Here, the focus shifts to how we filter that data for attribution. Accuracy begins not with scale, but with signal.

We prioritize information that points to real-world control—patterns of behavior, off-chain identities, and disclosures from trusted entities. This foundation ensures that every attribution starts with more than a transaction, it starts with context, credibility, and investigative value.

3.2 Validate with Structure, Not Assumptions

Once high-signal data enters our system, the next step is to determine whether it merits attribution. That process isn't automated guesswork—it's a structured pipeline designed to ensure every tag is explainable, consistent, and auditable.

We begin with automated rule checks that screen each entry for:



If the system flags uncertainty, the case moves to manual analyst review. Here, our team examines:

- On-chain behavior patterns (e.g. co-spend, withdrawal consolidation, routing structures)
- Source context (e.g. law enforcement disclosure, partner notes, open-source forums)
- Metadata: screenshots, URLs, TX hashes, or relevant case history
- Clustering logic, especially when tags intersect with existing infrastructure

If a case requires review, the analyst assesses whether the attribution is strong enough to move forward. If the evidence is clear, it’s approved and added to the platform. If the logic doesn’t meet our standards—due to weak sourcing, insufficient behavioral consistency, or conflicting signals—it’s rejected. In more complex cases, like when multiple entities appear to control overlapping infrastructure, the attribution is escalated for senior review before any action is taken.

This blend of automation and investigative review ensures that attributions aren’t just accurate in theory—they’re usable in practice. Whether for court, compliance, or cross-jurisdictional collaboration, they hold up.

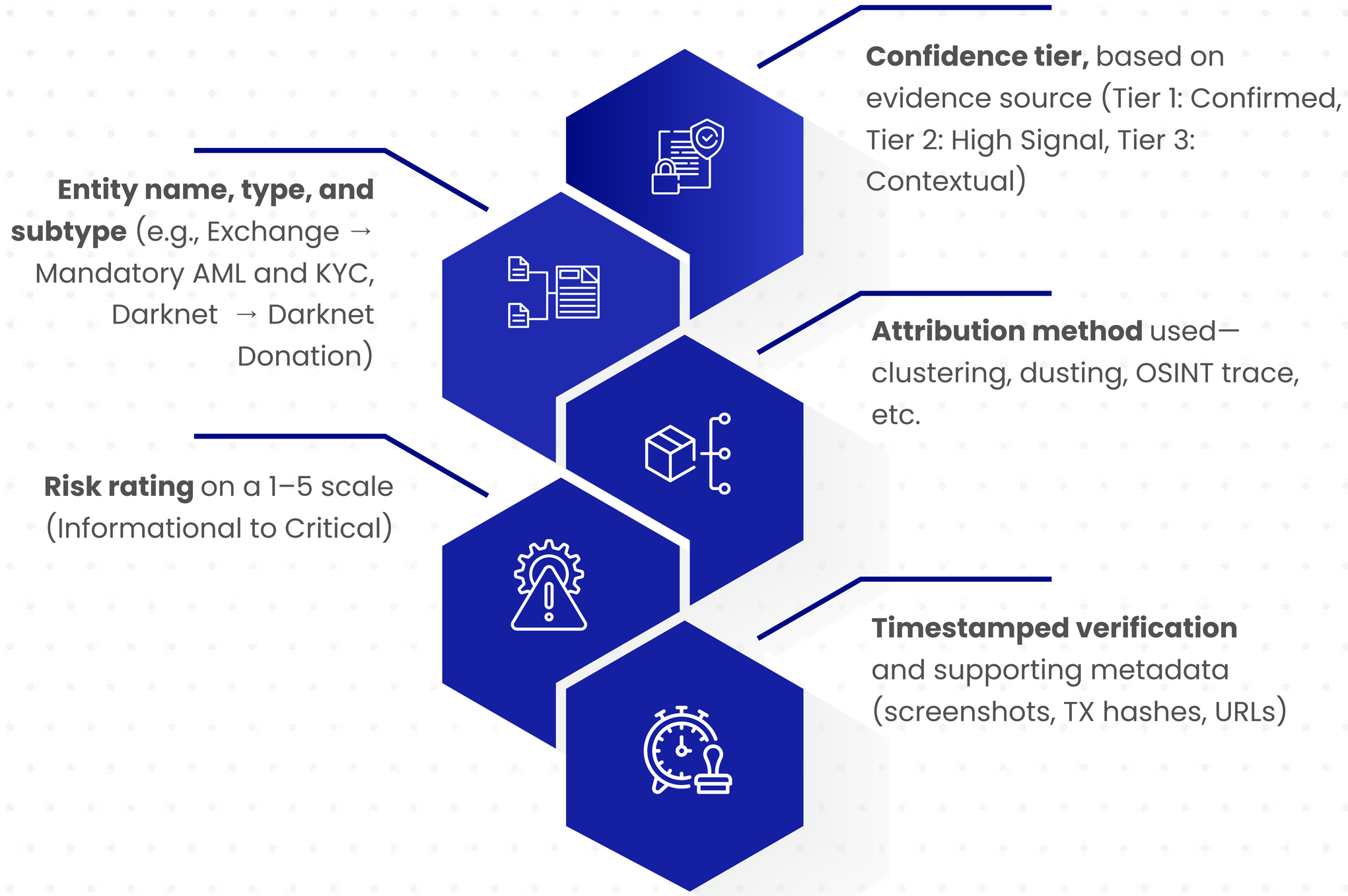
Note:

→ Automated checks are the first stage in a structured validation pipeline, with analyst review built in to ensure investigative integrity.

3.3 Structure Attribution for Actionability and Oversight

Once an attribution is validated, we make it usable—not just for tagging, but for investigations. That means translating raw input into structured, transparent metadata investigators can interpret, verify, and rely on.

Each address or cluster in our system is enriched with:



To ensure consistency and accountability, we run every attribution through a backend QA process which includes the following steps:

Entity name validation

Confirms that the tag includes a complete and valid entity name (e.g., “Binance” not left blank or abbreviated).

Naming format verification

Ensures the name follows our internal conventions so it’s readable and standardized across the platform.

Entity type and subtype consistency

Checks that the assigned type (e.g., Exchange → Custodial) matches the service's actual function on-chain.

Correct tagging of entity

Prevents misclassifications (e.g., tagging a smart contract address as a person or service).

Address tagging accuracy

Ensures address-level tags (e.g., "point" vs. "cluster") are correctly applied and not switched or duplicated.

Address format (regex) and duplication controls

Ensures the address follows valid chain format and hasn't already been tagged under a conflicting entity.

Cluster-to-entity relationship checks

Validates that clusters aren't split across multiple entities without explanation (e.g., Garantex and Hydra both claiming the same wallet).

Proof documentation and URL verification

Verifies that supporting evidence (e.g., screenshot, blog post, TX link) is valid, live, and tied to the correct entity or behavior.

Attribution logic is preserved across both address-level and cluster-level views, giving investigators visibility into the entire operational structure—whether it's a high-risk mixer, a backend wallet for a nested exchange, or a cold storage address tied to a darknet service.

Investigators don't just see who—it shows why and how the link was made. Every tag is traceable, every logic path auditable. This ensures that attribution in our system is more than a label. It's an accountable, evidence-based claim—fit for regulatory filings, casework, and cross-border collaboration.

3.4 Keeping Attribution Fresh

Attribution isn't a one-time action. It's a living layer of intelligence that must evolve as the blockchain does. Infrastructure changes, threat actors rotate wallets, and previously dormant addresses come back online. If attribution isn't monitored and updated, it becomes a liability. That's why Merkle Science builds data freshness into the core of our attribution system.

Attribution freshness refers to the last point in time when a tag was actively reviewed or confirmed, regardless of whether the entity itself changed. Every attribution entry includes a timestamped audit trail, allowing users to see exactly when it was last validated.

If attribution remains untouched for a defined period, or if the entity remains inactive beyond a certain threshold, our system automatically flags it for review. These flagged records are routed into a structured revalidation queue, where analysts assess whether:

- New OSINT or behavioral patterns warrant an update
- The entity should be reclassified or split
- The attribution remains accurate and can be reaffirmed as-is

And when our attribution logic evolves, whether through a new heuristic or updated validation rule, we conduct retrospective reviews. Past tags aren't left behind; they're re-evaluated and updated to reflect our latest understanding.

The goal is simple: every attribution in our system should be as accurate today as the day it was added. All updates are timestamped, versioned, and visible—so that investigators always work from the latest intelligence, not outdated guesses.

This approach ensures Merkle's attribution layer remains dynamic, traceable, and grounded in current data. By integrating attribution freshness checks into the platform's core logic, we reduce reliance on outdated assumptions and maintain confidence in the tags used across investigations, compliance actions, and enforcement workflows.

Handling Rebrands, Mergers, and Infrastructure Transitions

Attribution freshness does not require reattributing an address every time a business changes its name or ownership. When an entity rebrands or is acquired, Merkle Science preserves the existing attribution unless there is clear on-chain evidence that control of the address has changed.

The entity label is updated only if:

- Wallet infrastructure migrates to addresses clearly controlled by the new entity
- On-chain behavior aligns with new operational patterns
- Conflicting attribution signals emerge through OSINT or partner intelligence

This approach maintains continuity in historical records while avoiding premature or inaccurate updates. It also ensures that dormant or reused wallets are not misattributed based solely on organizational changes.

Attribution Update Triggers and Response Times

Trigger Event	Merkle Science Response	Typical Update Speed
OFAC or international sanctions	Address and cluster flagged + risk scored	Within minutes to hours
Law enforcement seizures or indictments	Attribution updated with source link	Same-day
Dormant wallet reactivation	Triggers revalidation workflow	Within 24 hours
Scam report or OSINT signal detected	New address or cluster tagged + reviewed	Within 24 hours
Heuristic behavior change (e.g. laundering tactics)	New rules tested + deployed	Within 30 days
Public infrastructure shift (e.g. new hot wallets)	Address surfaced via tracing logic	Real-time
Customer case annotation	Manually reviewed by analysts	Within 1 business day

3.5 Product Feedback Loop: Built to Learn from the Field

Even the best attribution systems can’t operate in a vacuum. That’s why we treat investigators not as passive users, but as active intelligence partners. The Merkle Science platform is designed to evolve—not just with the blockchain, but with the people using it to pursue criminal leads, seize assets, and dismantle illicit networks.

Through platform annotations, partner briefings, and forensic debriefs, we regularly incorporate frontline feedback into our attribution model. When users surface issues or insights, they don’t disappear into a black box—they become triggers for refinement.

Inputs that drive improvements include:

- Disputed tags flagged by law enforcement or compliance partners
- Confirmed seizures or court records that validate or reshape cluster logic
- Investigator notes submitted via the platform—capturing infrastructure reuse, tactic shifts, or false positives
- Field-detected patterns that prompt review of heuristic thresholds or clustering logic

All feedback enters a structured revision cycle. Our analyst team reviews the signal, correlates it with existing intelligence, and updates entity relationships or tagging behavior when warranted. In some cases, feedback directly results in new heuristics, changes to risk tiering, or the creation of specialized detection modules for known threat actors.

The system doesn't just learn from the chain—it learns from the people tracking it.

This closed-loop design ensures Merkle Science attribution is not just technically sound—but grounded in operational reality. When the field moves, we move with it.

Conclusion: Attribution That Holds Up

Attribution is only as useful as it is explainable. For investigators and enforcement teams, that means every tag must be traceable, auditable, and defensible under scrutiny.

In this report, we've outlined how Merkle Science builds attribution around structure, not speculation—from sourcing and clustering to validation, risk scoring, and real-time updates. Every address we tag is backed by structured, explainable intelligence designed for real-world investigations.

We also recognize that no system is perfect in isolation. That's why our platform is built to be self-correcting—refined continuously through real-world input from the people using it to trace risk, dismantle laundering systems, and build prosecutable cases.

When attribution becomes actionable, it becomes powerful. That's the standard we hold ourselves to.

If you're an investigator, analyst, or regulator looking to build trust in your attribution workflows, get in touch.
We'll walk you through real cases, explain our methodology, and explore how we can help you trace smarter, act faster, and stay ahead.

MERKLE SCIENCE

Contact us

contact@merklescience.com www.merklescience.com

Crypto Investigations Certification

Learn how address attributions support crypto investigations

Follow Us merklescience @MerkleScience merklescience Merkle Science