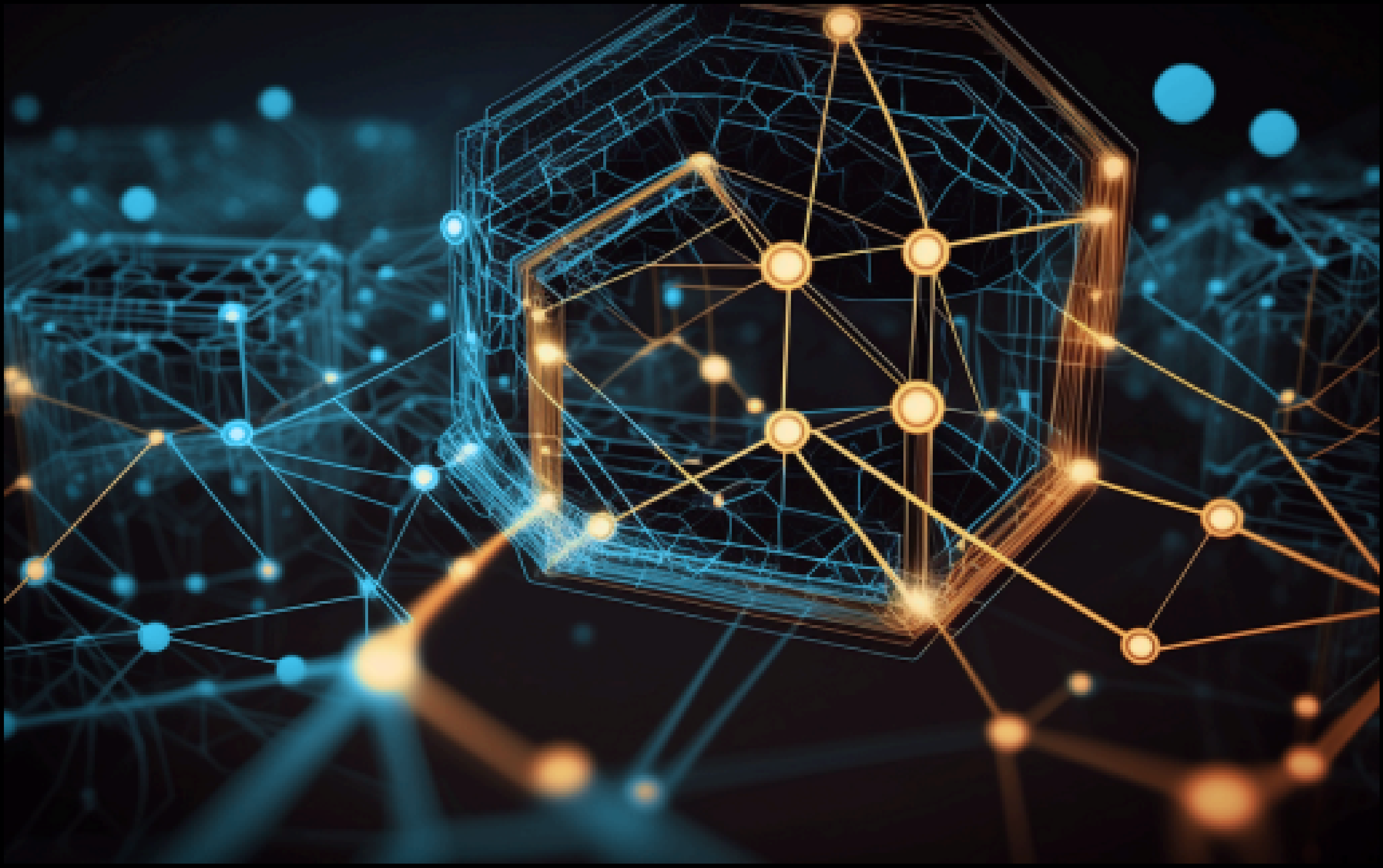
DIGITAL ASSETS LAWS & REGULATIONS

THE NEW PHRENOLOGY: HOW UNVERIFIED BLOCKCHAIN TRACING THREATENS THE INTEGRITY OF MODERN JUSTICE



WWW.LEXHOST.COM

Overview

The New Phrenology: How Unverified Blockchain Tracing Threatens the Integrity of Modern Justice

The advent of cryptocurrency heralded a revolution in finance, built upon the radical transparency of the public ledger. Yet, within this transparent system, mechanisms for privacy—such as Bitcoin mixing services—emerged, intended to maintain transactional confidentiality. The pursuit of criminals using these tools has, in turn, birthed a new and influential discipline – “blockchain tracing” or blockchain forensics.

Blockchain tracing has rapidly entered the legal lexicon, championed by investigative bodies and commercial forensic firms as an indispensable tool for identifying and prosecuting cybercriminals. However, the foundational methodologies underpinning this industry have largely bypassed the rigorous scientific scrutiny demanded of established forensic fields. We stand at a critical juncture where the legal system is increasingly relying on methodologies described by experts as “pseudo-science and conjecture,” risking a wave of wrongful convictions based on what critics have rightfully dubbed “junk science”.

This piece posits that without immediate judicial intervention to enforce scientific rigor and transparency, blockchain tracing risks following the shameful trajectory of debunked forensic methods like phrenology, handwriting analysis, and other practices that, for decades, polluted courtrooms and undermined justice.

The recent conviction and 12.5-year sentencing of Roman Sterlingov, the alleged founder of the Bitcoin Fog mixer, stands as a stark warning, encapsulating the systemic deficiencies and risks inherent in allowing untested commercial tools to determine criminal culpability.

The Illusion of Infallibility: Heuristics Masquerading as Forensic Science

The core danger of modern blockchain tracing lies in the fundamental difference between the immutable cryptographic truth of a public ledger and the fallible, proprietary assumptions used to link digital addresses to real-world entities. While a Bitcoin private key either works or it does not, providing cryptographic certainty in limited instances, other aspects of blockchain attribution are trivial to manipulate.

The industry's attribution methods rely almost entirely on "heuristics". A heuristic is essentially a "guess," or perhaps an "educated guess," used in the experimentation process. When claims are made that a heuristic is inherently reliable without independent proof or testing, that methodology must be called into question.

In USA v. Roman Sterlingov,1 694 F. Supp. 3d 1, 1 (D.D.C. 2024), the primary method in securing Roman's conviction relied on two heuristics, both of which demonstrate profound scientific deficiency when exposed to adversarial scrutiny:

The Flawed "Co-Spend" Heuristic

The primary methodology utilized by leading commercial blockchain tracing software, such as Chainalysis Reactor, is the "co-spending" heuristic. This heuristic suggests that when multiple inputs are spent in a single Bitcoin transaction, the sender must know the private signing key for each input, making it highly probable that all associated public keys are controlled by the same entity.

However, this assumption is demonstrably flawed and "unsubstantiated". Even before Bitcoin was created, it was shown that participants in a CoinJoin —a type of transaction designed for privacy—do not need to share private keys.

The assertion that co-spending implies common ownership is often taken entirely out of context and elevated to an axiom in a manner wholly inconsistent with accepted standards of forensic evidence.

Crucially, CoinJoins violate the co-spending heuristic. Chainalysis's own co-founder acknowledged that clustering techniques relying on co-spending "does not apply to CoinJoin transactions."2 Yet, the prosecution's entire attribution in Sterlingov appears to have relied on overlooking the overwhelming prevalence of CoinJoins.

The Epidemic of Obfuscation: The CoinJoin Problem

The integrity of the co-spend heuristic is decimated by the history of Bitcoin usage. CoinJoins—single Bitcoin transactions with multiple inputs and outputs —have existed since the blockchain's inception. Their use grew dramatically from 2009 through 2012.

In fact, Bitcoin transaction best practices, from the very inception, encouraged obfuscation and privacy. Satoshi Nakamoto himself advised using a new key pair for each transaction to prevent addresses from being linked to a common owner, explicitly highlighting the vulnerability of multi-input transactions to the co-spend heuristic.

Empirical analysis demonstrates that the prevalence of CoinJoins was "endemic rather than exceptional" during the early Bitcoin era relevant to Sterlingov. By late 2012, millions of CoinJoin transactions occurred monthly. The overwhelming prevalence of these privacy-enhancing transactions fundamentally undermines the reliability of heuristic-based attribution methods".

Furthermore, CoinJoins were highly effective in masking origins. Bitcoin Fog was at best responsible for only an “insignificant fraction” of all CoinJoins during its early operation, implying vast unknowns about other mixing activities during that critical period in Sterlingov.

This profound uncertainty about transaction origin makes applying heuristics extremely limited for transactions of average size. Without relevant external information to corroborate attribution, using such analysis to determine culpability is wholly disproportionate to its evidentiary value.

The Untested “Behavioral Heuristic”

The situation is compounded by the reliance on “behavioral heuristics,” techniques based on analyzing the “digital fingerprints” left behind by wallet software interaction. In Sterlingov, the prosecution claimed this method was “provable” and “very reliable”.

Yet, the same prosecution expert who made this claim also conceded that no academic studies or independent assessments of this heuristic’s integrity had ever been conducted. This claim is inherently problematic, especially since anyone can write their own transaction-signing program for Bitcoin, or even combine open-source wallet software into a “meta wallet” to randomly alter these “digital fingerprints”. Claims of reliability that have never been tested or scrutinized are fundamentally unscientific.

The Ghost of Forensics Past: Echoes of Debunked Science

The push to admit and rely upon unverified blockchain tracing methodologies into courtrooms draws strong, disturbing parallels with discredited forensic evidence of the past. History is
littered with examples of techniques, once deemed reliable and definitive, that were later exposed as having no scientific basis, leading to grave miscarriages of justice.

Consider the discredited science of phrenology — the practice of mapping character and intelligence based on skull measurements. Though wildly popular in the 19th century, it was ultimately proven to be baseless superstition, yet it held sway because it offered a seemingly scientific method to categorize and judge individuals.

Similarly, other forensic disciplines, such as handwriting analysis or microscopic hair comparison, were long presented to juries as definitive evidence of individual identification.

Today, courts recognize the lack of scientific foundation in such pattern-matching techniques. Handwriting analysis, for instance, is now treated, at best, as circumstantial evidence, rarely sufficient to prove guilt beyond a reasonable doubt because the error rates were never scientifically validated.

The risks posed by blockchain tracing are precisely those exposed by these historical failures: the presentation of subjective interpretations and assumptions as objective, quantitative fact.

The 2009 United States National Academy of Sciences (USNAS) Report, a landmark study to Congress, provides the necessary framework for scrutinizing these new digital methods. The USNAS Report highlighted that, with the exception of nuclear DNA analysis, “no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source”. Blockchain tracing, lacking independent corroboration, falls far short of providing a degree of certainty remotely close to DNA evidence.

The standards for digital evidence must align with the standards applied to physical evidence:

1. Transparency vs. Proprietary Secrecy: Early forensic techniques often relied on opaque methodologies or the subjective experience of individual practitioners. Today, robust forensic science demands transparency and peer review. However, in Sterlingov, the government relied on proprietary software and algorithms and Roman was denied full access to the closed source code and the proprietary heuristics of the tool used against him. Assuming a detection technique is reliable simply because its creators have not identified its vulnerabilities is fundamentally unscientific—a concept famously captured by “Schneier’s Law”.

2. Quantified Error Rates (False Positives/Negatives): Historically flawed forensics failed because they lacked quantifiable error rates. The USNAS Report emphasizes that acknowledging, studying, and managing errors is a core part of legal processes around forensic evidence. Blockchain analysis techniques, even in controlled studies, are often found to be inferior to existing, flawed forensic methods. Broader studies of wallet clustering generally achieved only 95% to 98% coverage, failing to classify 2% to 5% of samples, and never even attempted to measure error rates. When tested against the less-than-perfect standard of “contactless fingerprint matching,” which the National Institute of Standards and Technology cautioned users would encounter “difficulty with any forensic applications such as latent matching, or support of courtroom testimony,” blockchain techniques performed worse.3

3. The Fiduciary Relationship Precondition: Traditional equitable tracing rules —often necessary for dealing with mixed funds—historically required a fiduciary relationship. Although tracing itself is merely a process of identifying

assets, the stringent preconditions for applying equitable rules highlights the inherent complexity and legal history governing how courts handle non-identifiable assets. In contrast, blockchain tracing, often performed by commercial firms, leaps straight to definitive conclusions of attribution without establishing the legal or evidentiary foundation necessary to manage inevitable errors in mixtures, such as those created by CoinJoins.

The USNAS Report clearly states that lawyers and judges are often “unreasonably held to determine the reliability of forensic evidence, without being equipped with the necessary scientific methodologies to do so”. This context creates a fertile ground for the wholesale admission of digital junk science that lacks any foundation in principles of mathematics, statistics, and forensic science.

The Sterlingov Case: A Confluence of Flaws

The conviction of Roman Sterlingov for money laundering and operating Bitcoin Fog serves as the definitive example of how this dangerous reliance on unverified blockchain tracing manifests in the courtroom.

Roman's conviction hinged on proprietary attribution methodologies that fundamentally flawed, rendering the results entirely unreliable, and thus inadmissible. The issues range from fundamental technical defects to deeply troubling procedural failures regarding disclosure:

1. Falsely Claiming Infallibility

Perhaps the most egregious scientific overstep was the claim made by an expert witness for the prosecution regarding the testing procedures used by the prosecution's blockchain tracing tool. The prosecution expert testified in Sterlingov that, based on their dataset, they had found "no false positives".

This claim of a 0% error rate in real-world application is statistically absurd. Experts noted that Chainalysis' own later, and largely irrelevant, academic verification study (which took place years after the trial concluded) cited a non-zero error rate. Presenting such an exaggerated claim materially influences a jury, who are not equipped to make such determinations regarding statistical validity.

Furthermore, the prosecution entirely disregarded the significant false negative rate. Chainalysis Reactor failed to identify 20% of the Bitcoin Fog addresses that the government itself had identified. By comparison, the false negative rate for forensic latent fingerprint decisions was found to be only 7.5%. A 20% failure rate is a high error rate, definitively insufficient for determining the reliability of the underlying heuristics.

2. The Proprietary Methodology and Contradictory Logic

The analysis used to link Roman to the creation of Bitcoin Fog was predicated on linking his Mt. Gox account withdrawals to Bitcoin Fog's earliest transactions via a series of allegedly unusual transactions. This conclusion was based on the mistaken assumption that CoinJoins were rare, whereas, in reality, approximately one in three Bitcoin transactions during the period relevant in Sterlingov were CoinJoins, meaning the transactions were not "unusual" in any meaningful sense.

Furthermore, the application of the behavioral heuristic was inconsistent. The transactions in one set of transactions consistently left small amounts of unspent bitcoin behind, whereas another set of transactions (attributed to Bitcoin Fog) left no unspent amounts. Applying the prosecution's own behavioral heuristic in Sterlingov would imply different owners for these groups, yet the prosecution's expert reached a contrary conclusion. This inconsistency in applying the same heuristic to two identical sets of blockchain transactions highlights why these methodologies demand a higher degree of scientific rigor.

3. Undisclosed Errors and Due Process Violations

The evidentiary failings were compounded by procedural issues suggesting a possible violation of the defendant's due process rights under the Brady v. Maryland4 standard.

Roman was arrested in April 2021. The underlying statement of facts for his arrest warrant contained a typographical error in a key blockchain address, incorrectly identified as "1KWMex" when the correct prefix was "1KWMcx". This error is critical because blockchain addresses are like DNA sequencing—a difference of a single character indicates an entirely different individual.

The prosecution, during the trial, used corrected data, indicating that federal agents had performed the tracing anew using the correct address. This necessitates the existence of internal notes or communications documenting the initial error and its correction. Such documentation would directly undermine the expert witness's assertion of "flawless" procedures.

More importantly, the failure to disclose these records to the defense—information that could have been used to impeach the prosecution's witness regarding the reliability and infallibility of their methodology—raises severe questions about the integrity of the prosecution's evidence. The prosecution is not entitled to present evidence at trial unless it can demonstrate that the underlying methods followed required and accepted standards. By potentially concealing evidence of known procedural failures, the prosecution reinforced the dangerous illusion of infallibility.

The Immediate Need for Judicial Scrutiny

The current state of blockchain tracing jurisprudence threatens financial privacy and civil liberties. Americans should be free to transact on the blockchain without fear of being condemned based on unverified blockchain tracing methodologies. The distinction between finding investigative leads and presenting evidence of guilt is critical: a detective may accept a high error rate when searching for leads, but that low true positive rate is insufficient for court evidence.

The USNAS Report's constructive criticism holds profound relevance for this emerging field: "The judicial system is encumbered by, among other things, judges and lawyers who generally lack the scientific expertise necessary to comprehend and evaluate forensic evidence in an informed manner". When expert testimony exaggerates reliability, the legal checks and balances fail.

This court must demand that blockchain analysis evidence meet established scientific standards. This is not an extreme position; it is a necessity to avoid casting aspersions upon the legal and forensic science communities.

To protect the constitutional rights of defendants and safeguard the integrity of justice, judicial guidance must immediately require that blockchain analysis methodologies demonstrate:

1. Peer-Reviewed Reliability: Methodologies must be transparent, not proprietary, and subjected to external academic and peer review to establish their scientific foundation.

2. Statistical Validation: Error rates, including both "false positives" and "false negatives," must be quantified and disclosed.

3. Transparency of Assumptions: The reliance on heuristics, such as the co-spend and behavioral analysis, must be disclosed, and evidence must be presented as to why these assumptions are justifiable, especially given the historical prevalence of privacy-enhancing technologies like CoinJoins.

The failure to require such standards risks enshrining a new era of “junk science” in our courtrooms—a modern digital phrenology where defendants are convicted not by facts, but by the unsubstantiated “guesses” of opaque commercial software. The integrity of American technological leadership and the fundamental promise of due process depend on the judiciary demanding the same level of scientific rigor for blockchain evidence that has been established for every other form of technical evidence. We must fix this failure before a flurry of wrongful convictions permanently stains the digital age of justice.

Author Biography

Patrick Tan

ChainArgos

Patrick Tan is a lawyer, a former licensed fund manager and previously a commercial airline pilot with Singapore Airlines. He analyzes and manages legal risks related to crypto trading, speaks regularly at blockchain and investment events and his writing is featured in leading publications. He specializes in areas related to the overlap between securities law and its application to cryptocurrencies.

Firm Biography

ChainArgos

We’re delivering actionable blockchain intelligence.
Say “no” to pseudo-science and “yes” to blockchain intelligence you can count on for commerce, compliance, and crime-fighting.
ChainArgos was founded by passionate finance, technology, legal, and aviation professionals who have worked at some of the world’s largest financial institutions, technology companies, law firms and airlines, including Nomura, Barclays, Lehman Brothers, Morgan Stanley, Allen & Gledhill and Singapore Airlines.