Arkham: A Platform for Deanonymizing the Blockchain

July 4, 2023

Abstract

Blockchain transactions are inherently public, allowing any third party to view and investigate them. However, this raw transaction data is by default unprocessed and anonymous, and therefore useless. It needs to be analyzed and deanonymized to be rendered usable. Arkham is a crypto intelligence platform that systematically analyzes and deanonymizes blockchain transactions, showing users the people and companies behind blockchain activity, along with data and analytics about their behavior. To generate intelligence at scale, Arkham uses Ultra, a proprietary AI-powered algorithmic address matching engine. By systematically linking blockchain addresses to real-world entities, and providing aggregate data and analytics that give a complete picture of their behavior, Arkham reveals the activity driving digital asset markets, enabling market participants to outperform. This paper also introduces the Arkham Intel Exchange, a decentralized protocol for the buying and selling of data and labels designed to significantly increase liquidity in the intelligence market. By generating intelligence at scale and facilitating its spread, Arkham makes crypto markets more transparent, efficient, and just.

1 Introduction

Arkham is founded on seven theses about crypto and the role intelligence plays in it:

1. Deanonymization is destiny. In the early days of the internet everyone was pseudonymous. Now, people use their real identity online. MySpace to Facebook was a microcosm of this transition. The same process is happening with blockchain identities – consider the rise of ENS and of NFT profile pictures. Eventually, everyone’s blockchain identity will be linked to their real-world identity.

2. Access to crypto data will be decentralized. In traditional finance, data is hidden in black boxes guarded by established players such as brokers, exchanges, regulators, and banks [7] [11]. In crypto, raw transaction data is visible to everyone due to the public nature of blockchains, but in practice, it only becomes usable once processed, aggregated and analyzed by data platforms. Existing tools are expensive and limited to only a select few, but ultimately they will be disrupted by tools available to everyone.

3. Crypto is becoming a core part of the global financial system. Crypto will reach mass global adoption because it is a more efficient, reliable, and robust financial system. Once you use crypto, there’s no going back. Explaining to your bank on a wire form who you want to send your money to and why becomes intolerable after you’ve seamlessly made large transfers in a browser wallet. Any foundational technological shift encounters resistance and other difficulties, but overcomes them through a process of creative destruction [8]. It will be the same with crypto.

4. Crypto intelligence will be widely adopted. Depending on their production cost and use, technologies vary in how widely they are adopted. Only governments and major corporations own rockets, but billionaires use the same smartphones as everyone else. In a world where crypto is used by all, crypto intelligence tools are more like smartphones, because they are useful for anyone using crypto, from those tracking their stablecoin remittances back home, to multibillion dollar firms trading complex derivatives.

5. The crypto intelligence economy will be ≥$30 billion annually. As tools become more powerful and crypto reaches global adoption, crypto intelligence tools will serve as a real-time
map of the new financial system and will therefore be a utility for serious participants and observers. The market for conventional financial data is $30 billion annually, led by companies such as Bloomberg and Palantir [2]. Compared to traditional financial data, on-chain data allows for more useful analysis, especially when augmented by new AI-based analytics. These factors will make the market for crypto intelligence at least as large as its traditional predecessor.

6. **The future of crypto data is entity based.** Of the five Ws, Who comes first. To best understand crypto activity one must first and foremost know who is behind it. This is why Arkham provides intelligence at the entity level. Entity-level intelligence provides a general purpose tool that verticalizes the crypto data industries for trading, compliance, research, and portfolio tracking. It is the key to providing an all-in-one platform.

7. **Crypto intelligence will power self-regulation.** Self-regulation is necessary for any industry to be healthy, irrespective of the decisions and laws of external rulemakers. It involves founding institutions to safeguard trust and quality in an industry by monitoring, promoting and enforcing standards of conduct [10]. This concept is not new - for centuries, working communities such as guilds have machined standards and guidelines for practitioners to abide by. These institutions will develop naturally in crypto as users learn where to place their trust, reducing the need for external regulation and allowing the industry to better meet the challenges it faces. The investigation and verification of on-chain activity is the foundation of crypto self-regulation, and crypto intelligence tools will play a central role in this process [6].

In combination, these seven theses support a vision for a total crypto intelligence platform that provides deanonymized, entity-based intelligence to everyone. Such a platform needs to provide total coverage of the blockchain, service the full range of use cases, and be affordable for a normal person. It needs to systematically collect, aggregate, and attribute transaction data to form a complete picture of crypto activity.

## 2 The Intelligence Platform

There are three major challenges in the generation of crypto intelligence:

- **Collection:** Locating sources of relevant on-chain and off-chain data and gathering this data reliably and continuously.
- **Aggregation:** Combining this data into a single source of truth that can be queried to produce any information desired.
- **Attribution:** Using this data to assign real-world identities to addresses.

To solve these problems, Arkham has developed Ultra. Ultra is a proprietary AI system for blockchain data synthesis, which gathers on and off-chain data from a variety of sources and synthesizes them to form a single, scalable, amendable source of truth. Ultra allows Arkham to transform raw blockchain transaction information into a clear view of the entities using cryptocurrencies and their activities, including exchanges, trading firms, and individuals. At present, the Arkham platform contains over 350 million labels and 200,000 entity pages.

The core of the Arkham platform is the Profiler. The Profiler is a holistic view of an entity or address’s activity, including:

- Transaction history, which can be filtered and sorted by USD value, token, counterparty, and time
- Portfolio holdings
- Balance history
- Profit and loss
- Exchange usage
• Top counterparties

The Profiler is the core of the platform because it presents a suite of intelligence on an entity or address for a full picture of its activity. Additional features allow this entity information to be leveraged for deeper insight:

• Visualizer: Customizable network analysis of entities/addresses and counterparties, which can be filtered on any key dimension
• Private labeling: The ability to create private address labels and entities
• Alerts: Customizable transaction alerts based on size, entity, blockchain, and token
• Dashboards: Pages for custom groups of entities showing key information at a glance
• API: Direct access to Ultra for custom queries
• Archive: Historical point-in-time analysis of an entity’s crypto portfolio at any point in its history

2.1 Multi-Chain Integration

The scope of on-chain analysis is rarely limited to a particular chain, but analysis tools often are. In order to provide a complete intelligence picture, Arkham collects, aggregates, and attributes data across chains by default, employing Ultra to reconcile cross-chain data. However, Arkham is also able to distinguish data by chain and break it down, allowing for analysis of a specific chain or subset of chains as required.

2.2 Arkham Filters

The old paradigm of blockchain analytics lacked systematic transaction filtering. It was necessary to scroll through pages of transactions, or export to an external source, to find desired data. Arkham introduces systematic transaction filtering, allowing for transactions across the platform to be filtered and sorted according to key metrics, including time, token, USD value, and counterparty. Transaction filtering produces an order-of-magnitude increase in the efficiency of many analysis tasks. What previously required hours of scrolling through transaction pages can be done instantly with Arkham filters.

2.3 Profiler

Holistic information on entity or address activity is the core of the Arkham intelligence paradigm. The Profiler is the platform function for viewing a complete, real-time picture of what an entity is doing on-chain and what they have done in the past. Arkham’s user interface presents information in components called Units. Units display particular types of information. The Profiler contains four units:

• Portfolio: current and historical point-in-time holdings
• Historical Performance: balance history and historical profit and loss
• Counterparties: exchange usage and top address/entity counterparties
• Transactions: real-time-updating log of full transaction history

The Portfolio Unit displays the entity’s current on-chain portfolio, broken down by token and by chain. It also contains the Archive, which allows for comparison of portfolio holdings at any two points in time.

The Historical Performance Unit displays total portfolio balance in USD over time. It has two views—total balance history and profit and loss history. Profit and loss is calculated using as cost basis the USD value of tokens at the time of initial transfer to the entity.

The Counterparties Unit displays exchange usage by volume, separated into deposits and withdrawals, and shown over time. It also displays the top overall counterparties by volume.
The Transactions Unit allows for the most fine-grained analysis of entity activity. It displays the entity’s full transaction history, updated in real time. Crucially, as detailed in section 2.3, transactions can be sorted and filtered along multiple dimensions, allowing for key transactions to be surfaced even if buried deep in an entity’s history.

Together the Profiler Units provide a picture of entity activity at multiple levels of granularity, across multiple chains, through time. The completeness of this picture makes it useful for the full scope of analysis tasks. The Profiler views an entity from multiple perspectives and combines these views into one complete picture.

2.4 Visualizer
The Visualizer enables network analysis of entity relationships, generating a network graph for a selected set of addresses or entities. Visualizer networks comprise nodes representing entities, linked by transactions. They are fully customizable using Arkham filters to display only transactions of interest.

Whereas the Profiler provides a complete picture of entity behavior, the Visualizer is suited for analysis of entity relationships. It allows for relationships among entities to be grasped intuitively, generating insights that would otherwise be difficult or impossible to obtain.

2.5 Alerts
Alerts are for monitoring and responding to activity of interest in real time. They notify Arkham users, via the platform of their choice, of transactions that meet custom criteria set using Arkham filters, instantly catching transactions of interest in the constant swarm of network activity.

2.6 Dashboard
The Dashboard allows users to collect entities and addresses of interest and build custom data feeds. A list of tracked entities and their recent transactions are displayed on the dashboard’s home page. Users can create custom dashboards for entities and data of interest, e.g. “market makers”, “ETH whales”, “OTC desks”, etc. These dashboards comprise custom units of various types, such as holdings, balance history, and transactions for one or multiple entities. Users can easily share dashboards with others via URL.

3 Intel-to-Earn & the Arkham Intel Exchange
The Arkham Intel Exchange connects people who want to buy and sell crypto intelligence. It is a decentralized intelligence economy where anyone can trade the native currency—ARKM—for information, such as entity labels, hacker tracing, and curated data feeds. Buyers request intelligence by placing bounties. Sellers offer intelligence by conducting auctions. Both bounties and auctions are conducted via audited smart contracts, with no centralized entity ever taking custody of funds.

By creating a market for intelligence, The Arkham Intel Exchange and accompanying DATA program (see section 3.6) allow people who generate intelligence to capture some of the value of their work—what we call ”intel-to-earn”. Intel-to-earn incentivizes the production of intelligence as a public good.

3.1 Exchange Concept
The economy for blockchain intelligence is fragmented and inefficient. There is significant demand for intelligence from traders, journalists, researchers, and other observers, and there is significant supply generated by the thriving community of on-chain sleuths, but there is nowhere for these two sides of the market to meet. The Arkham Intel Exchange solves this problem by providing a platform for buyers and sellers to interact and exchange intelligence. By enabling the free flow of intelligence, the exchange increases the efficiency of the intelligence market and consequently of the crypto market as a whole.

The Arkham Intel Exchange is the first decentralized network to power, incentivize, and reward on-chain analysis at scale. It allows Arkham and its users to source intelligence from the thriving
community of on-chain sleuths, which has been experiencing steady growth for the past two years and rapidly increased in size throughout 2022. These sleuths currently earn nothing for their work and technical expertise [9], but this will change with the Arkham Intel Exchange and the advent of intel-to-earn.

Buyers seeking particular intelligence place bounties by staking ARKM. The staked assets are the initial bounty. Once the initial bounty is posted to the Arkham Intel Exchange, then anyone can add to it by staking the same amount or more ARKM, thereby increasing the value of the bounty and of the incentive for bounty hunters. Once the initial bounty is posted to the Arkham Intel Exchange, then anyone can add to the bounty by staking the same amount or more ARKM. Anyone can compete for the bounty by making a submission, which then goes through the intel verification process. The first verified submission receives the staked bounty.

Sellers can also set up auctions for intelligence. If a seller has information they believe is valuable to the market, they can submit it for verification. Once verified, the intelligence is put up for auction. Sellers can set standard auction parameters including end-date and a buy-now price. Then interested parties bid for access to the intelligence.

Bounty stakers and auction winners receive purchased intel exclusively for 90 days. Afterwards, it may be propagated to the broader Arkham platform for everyone to use. All ARKM staking and payout is controlled by audited smart contracts, without Arkham ever taking custody of the ARKM. To support the network, Arkham takes a 2.5% maker fee on submitted bounties and auction payouts and a 5% taker fee on bounty payouts and successful auction bids.

Intel submitted in bounties and auctions is reviewed by the Arkham Foundation, which either approves or rejects it. The Foundation’s response is validated on-chain using a Chainlink Decentralized Oracle Network (DON) [1] [4], which serves as the oracle for the bounty and auction smart contracts. The use of a DON secures and decentralizes the intel verification process, mitigating the risk of false verification. It also lays the groundwork for complete decentralization of the intelligence review process, described below in section 3.5 on the intelligence oracle problem.

These mechanisms produce an ecosystem of overlapping market roles:

- **Bounty posters**, who post bounties for particular sought-after intelligence
- **Bounty hunters**, who hunt for intelligence to claim posted bounties
- **Intel merchants**, who auction-off intelligence they gather themselves or receive from others
- **Intel collectors**, who collect intelligence from auctions rather than from targeted bounties

ARKM is the currency of the Arkham Intel Exchange.
3.2 Bounties

Buyers solicit intelligence by posting ARKM bounties on the exchange. To post the bounty, they must write the specifications of the information they seek, then lock in the bounty contract the amount of ARKM they wish to post as a reward. Anyone seeking the same intelligence can join the bounty by locking an equivalent amount of ARKM, which is added to the total reward.

If the original poster of a bounty wants to remain the only intel recipient after others have joined, they can do so by buying out the joiner stake. For example, if the original poster placed a 1,000 ARKM bounty, and then two others join the bounty with 1,000 ARKM each, then the original poster can buy them out of the bounty by staking an additional 2,000 ARKM. The joiners receive their ARKM back, while the total value of the bounty remains at 3,000 ARKM, reflecting the increased demand for that intelligence. Joiners who have been bought out can buy back in by staking the new total bounty amount, in this case 3,000 ARKM.

Bounty hunters claim the bounty by submitting intelligence on the exchange. To prevent spam, bounty hunters must stake 10 ARKM in the smart contract, which they lose if a submission is rejected.

If the Arkham Foundation approves a submission, the exchange DON registers this response and reports it to the bounty smart contract, which then starts the 15-day unlock timer. After this period has elapsed, the bounty hunter is able to withdraw all ARKM in the bounty contract, less the exchange fee, to the address used to stake submission ARKM. They may withdraw before the end of the 15-day lockup period for a 10% fee. If, on the other hand, the submission is rejected, the submission ARKM remains in the smart contract to be added to the bounty.

3.3 Auctions

Sellers put intelligence up for sale by conducting auctions. Arkham Intel Exchange auctions will follow the same paradigm as auctions on NFT marketplaces, where sellers can set a purchase price, minimum bid, and length of auction. Once the seller submits an auction post, the intelligence is verified by Arkham before the auction begins. To prevent spam, auction holders stake 10 ARKM with their
Figure 3: Bounty Structure

Figure 4: Auction Structure
submission, which is lost if the submission is rejected. Auctions, like bounties, have a 15 day lock-up before a winning bid can be withdrawn from the auction smart contract, but auctioneers can withdraw early for a 10% fee.

3.4 Implementation

Bounties and auctions can be conducted for any kind of intelligence, not just labels and entity attributes. For example, victims of exploits might pool resources to offer large bounties for intelligence on the exploiter, or a sleuth might auction a particular method for gathering labels.

Intelligence bought and sold on the Arkham Intel Exchange is held exclusively by the acquiring counterparties for a period of 90 days, after which it will be propagated to Arkham for all users to see. Time-bound exclusivity is necessary so that Arkham is not constrained from putting requested intelligence on the platform, and to contribute to the ultimate decentralization of intelligence for the community. If exclusivity were unlimited, then Arkham would not be able to add the same intelligence to the platform even if acquired independently. Indirect intel verification as described in section 3.5 would allow for unlimited exclusivity.

The user interface for interacting with the Arkham Intel Exchange will be deployed on the Arkham Platform. The UI will display a searchable list of the existing intel bounties and auctions on the exchange, and provide a portal for posting new bounties and auctions. Bounty and auction listings display a description of the intel sought or offered, with options to claim or bid.

Bounty posters enter a description of the intelligence sought, the time limit for the bounty, and the bounty size. They then use their browser wallet to stake the bounty ARKM in the bounty smart contract. Once the transaction clears, the bounty is posted to the exchange.

Auction holders enter a description of the intelligence sought, the time limit for the auction, and the bounty size. They then use their browser wallet to stake the bounty ARKM in the bounty smart contract. Once the transaction clears, the bounty is posted to the exchange.

Auction holders enter a description of the intelligence sought, the time limit for the auction, and submit the intelligence to be sold for review. Once their staking transaction clears, the submission is reviewed and, if approved, posted to the exchange under the conditions set by the auctioneer.

3.5 The Intelligence Oracle Problem

Intelligence markets are uniquely challenging in that they require indirect verification. A buyer can’t directly assess the intelligence prior to a sale, because then they would possess it, and there would be nothing left to sell.

To get around this problem, buyers often use proxies for intelligence quality, such as the seller’s reputation, but these are error-prone and invariably constrict possible sources of intelligence. Going off of reputation, for instance, leaves untapped the well of crypto intelligence sourced by amateur sleuths who have yet to establish their reputation, or by those wishing to remain anonymous.

Arkham solves this problem by directly verifying intelligence, so exchange participants know that intelligence on auction or successfully submitted for a bounty has been vetted by the Arkham team using proprietary technology.

Eventually, intelligence on the Arkham Intel Exchange will be verified in a decentralized manner. This requires solving the oracle problem for crypto intelligence. Arkham is soliciting grant proposals for work to solve this problem.

The intelligence oracle problem is a level of difficulty above the general oracle problem because of the requirement for indirect verification. It is not as simple as reliably retrieving the correct piece of off-chain information. The soundness of the information must be verified in a reliable decentralized way without direct access by validators.

Existing solutions to the general oracle problem, such as Chainlink DONs [1] [4], provide a potential starting point for work on the intelligence oracle problem. However, DONs currently only address the problem of oracle decentralization, not indirect verification.

Promising avenues for solving indirect verification include zero-knowledge protocols [5] [3], which are designed for this purpose. Thus far they have primarily been employed to verify private key validity, which is less challenging than indirect intelligence verification because it is less subjective. Intelligence verification always involves some uncertainty and ambiguity, making it especially well-suited to direct human analysis. However, there are existing methods for zero-knowledge assessment in ambiguous domains that present promising research avenues.

In sketch, a zero-knowledge decentralized solution to the intelligence oracle problem might involve a decentralized network of intelligence analyst nodes which verify intelligence that has been concealed.
using a zero-knowledge protocol. Compensation for the work of these intelligence validator nodes could be built into the Arkham Intel Exchange protocol through allocation of a portion of staked bounties. A benefit of deriving validator compensation from staked bounties would be that larger bounties would be prioritized for validation, providing an additional incentive to increase bounty sizes.

The Arkham Intel Exchange will employ a DON in the intel verification process. The DON serves to decentralize and secure the intel verification oracle for bounty and auction smart contracts, but the verification itself remains direct. When an indirect verification method is devised, the DON can be easily modified to refer to this new method.

### 3.6 The DATA Program

Arkham’s Decentralized AI Training Accelerator (DATA) Program is a complementary intel-to-earn system to the exchange. In the DATA program, sleuths earn ARKM by submitting intelligence for training Ultra and augmenting the data on the platform. In addition to specific labels and entity attributions, data sources and other leads can be submitted. The Arkham Foundation reviews submissions and, if approved, there is an ARKM payout in accordance with the DATA price list. The direct submission system provides sleuths with a revenue stream for their work, while also creating a decentralized community of intelligence analysts for the platform.

Direct submissions are a means for Arkham to source intelligence and training data directly from the community, and for sleuths to earn ARKM for intelligence without the uncertainty of auctions and the restricting specificity of bounties. With direct submissions, rewards are given for any approved intelligence submission in accordance with a preset price list, shown in Table 1. There are three types of submissions: address labels or disputes, entity submissions, and high priority intel.

High priority intel includes a range of non-address, non-entity submissions, such as data sources that generate many labels or intelligence on an urgent and high-profile case. Sleuths participate in the DATA program by submitting intelligence directly on the platform using a simple form in which they present their findings and justify their validity. Approved submissions receive the designated payout.

### 4 ARKM Incentives

In addition to its main function as the currency of the intel-to-earn economy, ARKM powers a system of incentives designed to generate a positive feedback loop of platform adoption, driving stable, long-term growth.

This incentive system has two parts:

- **ARKM Rewards**: Rewards for actions beneficial to the Arkham ecosystem
- **ARKM Discounts**: Platform discounts of up to 60% for paying in or holding ARKM

Intel-to-earn forms the core of ARKM’s unique utility as the first crypto intelligence token. The system of incentives formed by ARKM discounts and rewards supports this core utility to complete an Arkham ecosystem that will stably flourish as the global center for blockchain intelligence.

#### 4.1 ARKM Rewards

The first component of the ARKM incentive system is the Rewards Program, in which users receive ARKM for actions beneficial to the Arkham community. Actions for which rewards will be distributed will vary depending on the state of the ecosystem at a given time, but may include user referrals and the production of research and analysis.

<table>
<thead>
<tr>
<th>Intelligence Type</th>
<th>ARKM Reward</th>
</tr>
</thead>
<tbody>
<tr>
<td>Address Label/Dispute</td>
<td>200</td>
</tr>
<tr>
<td>New Entity Submission</td>
<td>400</td>
</tr>
<tr>
<td>High Priority Intel</td>
<td>≥1,000</td>
</tr>
</tbody>
</table>

Table 1: DATA Price List
ARKM Holdings | 30 Days | 60 Days | ≥ 90 Days
---|---|---|---
$100 | 2% | 4% | 5%
$1,000 | 5% | 7% | 10%
$10,000 | 10% | 15% | 20%
$100,000 | 20% | 25% | 30%
$1,000,000 | 30% | 40% | 50%

Table 2: Holding Discounts

4.1.1 Grants
ARKM grants will support projects that contribute to the ecosystem, especially work to solve the intelligence oracle problem towards achieving full exchange decentralization.

4.1.2 Ecosystem Partner Rewards
The vast majority of on-chain sleuths focus on Ethereum. Other blockchains will be able to offer rewards to Arkham users for contributing on-chain analysis and other content to their communities. Rewards may be in partner native tokens or in ARKM.

4.2 ARKM Discounts
ARKM Discounts are the other main aspect of the ARKM incentive system. Users earn discounts on the Arkham platform by using ARKM in two ways:

- **Pay-in-ARKM:** 20% discount on platform payments made in ARKM.
- **Holding Discount:** Up to a 50% discount for holding ARKM, varying by holding size and duration in accordance with the table above. In order to prevent abuse of the system, for example via short-term borrowing, users must lock ARKM for at least 30 days in order to receive holding discounts.

These discounts can be combined, meaning that a user can pay in ARKM and receive the 20% discount, while also retaining ARKM holdings to receive the holding discount as well, for a maximum discount of 60%. Pay-in-ARKM discounts are implemented using a browser wallet. Holding discounts are implemented via simple audited smart contract from which users can withdraw tokens at any point after the locking period has elapsed.

5 Conclusion

Compared to traditional financial systems, blockchain-based technology enables more complete and precise analysis of financial activity. The greater power of crypto financial data compared to traditional data confers a significant advantage upon those who can use it, rendering crypto intelligence tools essential for serious market participants. Their advent will transform the financial industry. This transformation, and the recognition of it, has only just begun. It will culminate in the development and adoption of an all-in-one tool that provides total crypto intelligence. Arkham is designed to be that tool.

References


