

Microscope Protocol Taxonomy

feedback.microscope@gmail.com

Microscopeprotocol.xyz

September 2023

Introduction

Current protocol design covers blockchain metadata at the address level, including key information of the addresses' *name*, *source*, *entity*, *chain*, and *category*. Apart from the blockchain address name, our data model defines a limited set of values for the rest of the schemas.

The taxonomy document provides additional information of these schemas, including the limited set of values defined in the data model and their detailed descriptions. The objective is to help our users to better understand the metadata dataset.

Source

There are many different ways to generate blockchain metadata and the 'Source' value provides information about how the given data entry was collected. We have five source values set at the data model and they are:

Ground Truth. This method relies on conducting real life transactions on chain to identify blockchain addresses that belong to a specific entity - it is especially useful for identifying addresses that are owned by crypto platforms and applications. For example, if we want to find out which addresses are owned by Exchange A on Ethereum, we can deposit a small amount of ETH or ERC20 token to Exchange A.

Heuristic. This method clusterize blockchain addresses using data models built from publicly available information. The widely adopted approach is the rule-based heuristic using historical transaction data on chains - i.e. labelling addresses based on transaction patterns. The two most common rule-based heuristics are Commonspend heuristic for UTXO chains and the Core-address heuristics for the account-based chains.

Machine Learning. This method utilised machine learning algorithms with on/off-chain data to train the address clusterisation models.

Research. This method relies on labellers to conduct thorough research over publicly available information. It is especially useful when there are significant events happening in the space - either good or bad.

External. This methods collects metadata from other open sources such as Etherscan.

Chain

The blockchain or the network where the given address is resided. The protocol has a fixed list of values defined at the data model but new chain values can be added to accommodate metadata from new blockchain networks.

Entity

The entity-level owner of a given address. For example, Uniswap addresses may be named slightly differently as they are used for different purposes - e.g. Uniswap V2: Myh, Uniswap V3 SUDT-BSFC, etc. These addresses will share the same entity name as 'Uniswap'.

The protocol has a fixed list of values defined at the data model but new entity values can be added to accommodate metadata from new blockchain networks.

Category

The 'Category' value provides the additional information about the purpose or the sector of the address. For example, weather the address belongs to a centralised exchange, or whether the address is used by a DeFi project.

The protocol has a fixed list of values defined at the data model (as shown below) but new category values can be added to accommodate metadata from new blockchain networks.

Category	Description
Cold Storage	An offline wallet used to securely store large amounts of cryptocurrency for extended periods of time.
Hot wallet	An online wallet that offers easy and quick access to cryptocurrency funds for everyday use.
Deprecated	A smart contract that is no longer in use and has been deprecated by the protocol or development team.
Multisig	An address that requires approval from multiple parties to access cryptocurrency funds, providing an additional layer of security. Commonly used by exchanges and institutional investors.
Vault	A smart contract-based system that automatically invests cryptocurrency funds to generate yield for users, commonly used in decentralized finance (DeFi) applications and managed by the protocol or community.
Vulnerable	An address whose private keys have been compromised or are publicly available, exposing it to the risk of unauthorized access or theft.
Smart Contract	An address used by developers or teams to deploy smart contracts on a blockchain for decentralized applications or other projects.
NFT	An address controlled by an NFT marketplace that facilitates the buying, selling, or trading of non-fungible tokens.

DEX	An address or wallet used in decentralized exchanges to facilitate peer-to-peer trading of cryptocurrencies without a central authority.
CEX	A centralized exchange (CEX) that allows users to trade cryptocurrencies in exchange for fiat or other cryptocurrencies.
ATM	Related to a cryptocurrency ATM that allows users to buy or sell cryptocurrencies using cash or debit/credit cards.
Auction	Related to an auction that can be held by a variety of entities, including law enforcement agencies (LEA).
Gambling	Related to a gambling website that allows users to place bets on various games using cryptocurrencies.
Gaming	Related to a gaming platform or game that allows users to buy, sell or trade digital assets or in-game items.
LE (Law Enforcement)	An address or wallet certified to belong to a law enforcement agency (LE).
Lending	Related to lending or borrowing activities in the crypto space, such as those belonging to lending platforms or peer-to-peer lending services.
Merchant	Addresses or wallets that are associated with merchants who accept payment in cryptocurrencies.
Mixer	Addresses or wallets that are associated with crypto mixers or tumblers, which are services that mix multiple transactions together in order to obfuscate the origin of the funds.
Payment Processor	Related to a business that facilitates transactions between buyers and sellers using cryptocurrency as payment.
Business or Services, Other	Related to a business or service that operates within the crypto ecosystem but does not fit into any of the other categories.
Token Team	An address or wallet that is controlled by the team responsible for creating and maintaining a particular cryptocurrency token.
Weapons	An address or wallet that is involved in the sale or exchange of weapons using cryptocurrency.
Marketplace Aggregator	Related to a business or service that consolidates data from multiple decentralized exchanges or NFT markets in order to provide users with a comprehensive view of the market.
L2	An address or wallet associated with a layer 2 scaling solution built on top of the Ethereum network.
Insurance	Related to A business or service that provides insurance coverage for cryptocurrency assets stored on the blockchain.
Asset Management	Related to A business or service that manages cryptocurrency assets on behalf of clients.
Streaming	Related to A business or service that provides streaming media content using cryptocurrency payments.
Infra	Related to A business or service that provides infrastructure services to support the operation of blockchain networks.
Oracle	Related to A service that provides off-chain data to smart contracts. Oracles allow smart contracts to interact with external data sources, making them more versatile and useful.
DAO	A DAO is an organization that is run by rules encoded as computer programs called smart contracts.
Yield	A yield service allows users to earn interest on their cryptocurrency deposits.
Wallet	A wallet is a software or hardware device used to store cryptocurrency private keys and sign transactions.
Fund	A fund is an entity that pools capital from multiple investors and uses it to invest in a variety of assets, including cryptocurrencies.
Derivatives	A derivatives entity is an organization that specializes in trading cryptocurrency derivatives, such as futures, options, and swaps.

E-Commerce	An e-commerce entity is an organization that uses cryptocurrency as a payment method for goods and services.
Identity	An identity entity is an organization that provides on-chain identity-related products, such as decentralized identifiers (DIDs) and verifiable credentials.
Mining Pool	An organization that brings together multiple miners to collectively mine cryptocurrency.
Liquid Staking	Companies that allow users to stake their PoS tokens and receive liquid staking derivatives in exchange.
Payments	Associated with companies providing on-chain payment processing services, allowing users to send and receive cryptocurrency payments on the blockchain.
Sidechain	Associated with sidechain solutions built on top of the Ethereum blockchain, allowing developers to deploy decentralized applications with increased throughput and scalability.
Router	Associated with router services that allow users to swap tokens across multiple liquidity pools.
Coin Swapper	Associated with services that allow users to swap one cryptocurrency for another without requiring users to deposit funds into the service wallet first.
Sport	Associated with sport-related products or services, including fantasy sports, sports betting, and collectibles.
Privacy (not Mixer)	Associated with products or services that provide privacy solutions on the blockchain.
Validator	Associated with validators that help secure Proof-of-Stake (PoS) blockchain networks.
Charity	Controlled by an organization that accepts crypto donations for charitable causes
Donation	Controlled by an organization that accepts crypto donations for any purpose
Deployer	Addresses that has been used to deploy smart contracts
Company Funds	Addresses managed by companies, projects or foundations, which are usually used for treasury management, paying salaries or other expenses
Eth 2 staker	Addresses that have staked Ether at the Ethereum 2.0 Beacon Chain to help secure the network and earn rewards
Yield Farming	Addresses used for yield farming strategies where users can deposit cryptocurrencies to earn rewards and/or fees on a decentralized platform
ICO	Addresses used to receive funds during an initial coin offering (ICO) or other crowdfunding mechanisms
Stablecoin	Addresses used for stablecoin projects, including centralized and decentralized stablecoins
Staking	Addresses used for staking, where users can lock up their cryptocurrency to help secure a network and earn rewards
Token Sale	Addresses used during a token sale where users can buy tokens at an initial price, including initial coin offerings (ICO) and initial dex offerings (IDO)
Defi	Addresses used for decentralized finance (DeFi) projects, including decentralized exchanges, yield farming, lending, and more
Advertising	Addresses used for advertising-related activities in the crypto environment.
Energy	Addresses used by entities that are bringing energy-related products into the blockchain.
FIAT	Addresses controlled by businesses that allow customers to trade fiat currencies.
KYC	Used for exchanges or other services that require customers to provide a valid ID document to trade fiat-to-crypto or crypto-to-crypto.
NO KYC	Used for exchanges or other services that do not require customers to provide a valid ID document to trade fiat-to-crypto or crypto-to-crypto.
Darknet	Addresses belonging to services hosted on the dark web, or being involved in dark web transactions.

Money Laundering	Addresses certified belonging to money laundering activities.
Malware	Addresses belonging to malware campaigns.
Ransom	Addresses that have been identified as being involved in ransomware attacks.
Scam	Addresses that have been identified as being involved in fraudulent activities or scams.
Terrorism	Addresses that have been identified as being involved in financing terrorism.
Theft	Addresses that have been identified as being involved in theft or hacking of cryptocurrencies.
Sanctioned	Addresses that have been sanctioned by some government in the world (excluding the US).
Phishing	Addresses related to phishing scams
Ponzi	Addresses related to Ponzi schemes
Spam	Addresses related to spamming
Fake KYC	Addresses related to fake KYC procedures
Malicious Mining Activities	Addresses used for mining cryptocurrencies with malware or botnets
Financial Crime, Other	Addresses related to financial crimes such as money laundering, fraud, and tax evasion
Cybercrime	Addresses related to cybercrimes such as hacking, data theft, and ransomware
Honeypot	Addresses related to fake tokens or ICOs that are designed to scam investors or traders
Black Mail Activities	Addresses that threaten to reveal sensitive information or personal data in exchange for ransom payments
Constrained by Service	Addresses that have been restricted or blocked from using centralized or decentralized exchanges due to suspicious activities or high risk status
High Risk, Other	Addresses that are associated with suspicious, fraudulent, or illegal activities that are not covered by other categories
Dapp, Other	Addresses used by decentralized applications (dApps) to interact with smart contracts and blockchain networks
Bridging	Addresses used by platforms or projects that provide interoperability solutions to connect different blockchain networks
Multichain	Addresses used by platforms or projects that enable communication and transfer of assets between multiple blockchain networks
Factory Contract	Smart contracts used to create and deploy multiple instances of other smart contracts
Fraud Proof	Smart contracts used to prove fraudulent activities or invalid state transitions on layer 2 protocols
Genesis	Genesis addresses used to mint/burn
MEV Bot	Addresses used by MEV bots for transaction extraction and prioritization
Proxy	Contract that acts as a proxy to another contract and is used to upgrade the logic of the target contract without changing its address
Exchange	Addresses used by centralized or decentralized exchanges to interact with smart contracts and blockchain networks