Lukka®

AUGUST 2026

WHITEPAPER

Compatible by Design

Extending the Security Master for Tokenized Real-World Assets

Compatible by Design

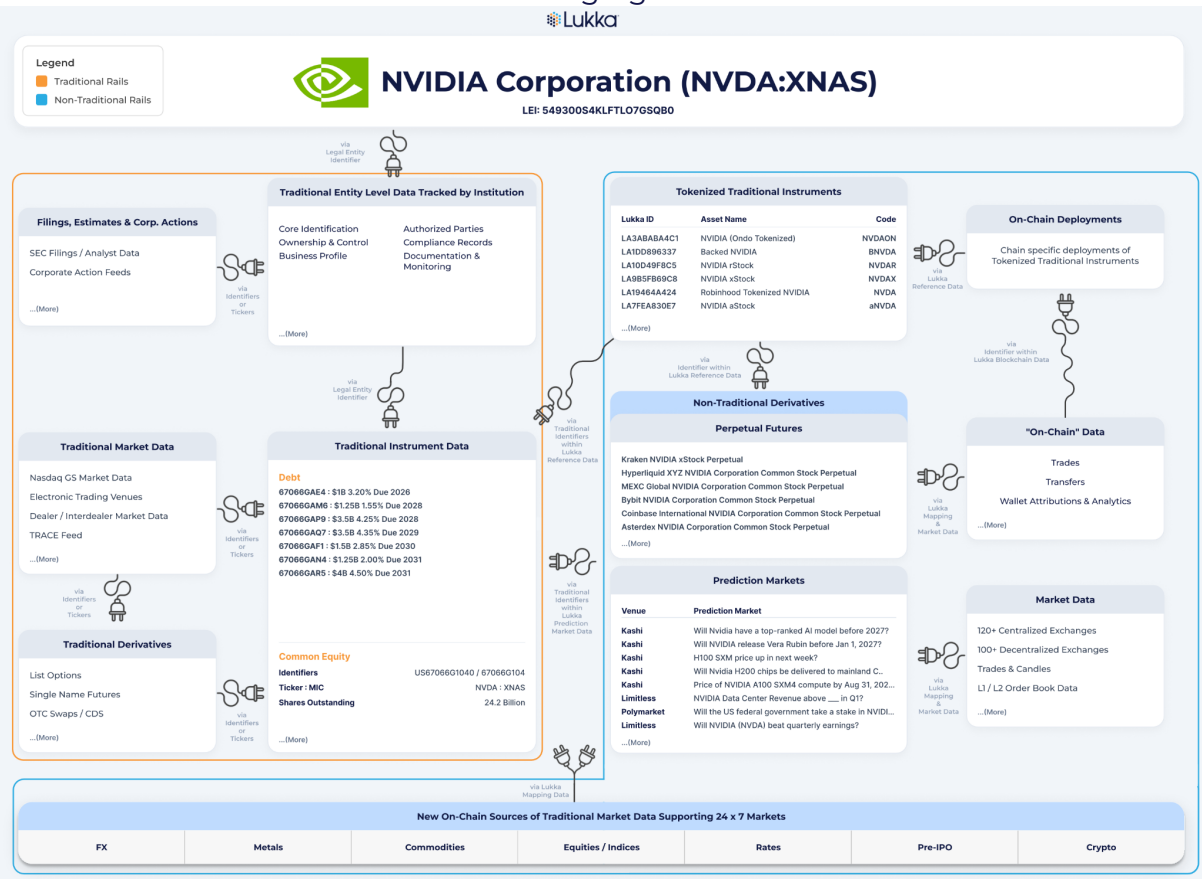
Extending the Security Master for Tokenized Real-World Assets

Executive summary

Tokenized real-world assets have moved from pilot to production. The instruments arriving on institutional books are no longer experiments run out of a digital assets desk: they are share classes of regulated funds, deposit tokens, money market vehicles, and collateral instruments,that are subscribed through normal channels and subject to the same fiduciary, audit, and reporting obligations as everything else in the book.

The Convergence Map

Where TradFi and DeFi meet: mapped across every dataset, on-chain and off.
Markets are converging. Data should too.



What has not moved is the reference data layer underneath them.

Security master systems handle multiplicity well along the axes they were built for. A cross-listed equity carries several identifiers, trades on multiple venues under different settlement conventions, and is priced through a defined source hierarchy. All of that is standard. What they were not built for is a class of attributes that had no variance in traditional markets and therefore never became fields. Nobody needed to record whether an issuer could unilaterally freeze a holder's position, or where the authoritative register of holders actually sat, or whether a holding in one venue was fungible with the same instrument in another, because those answers were structurally fixed. Tokenization makes every one of them a variable.

A tokenized fund can exhibit all of this at once. A single legal share class may be represented across multiple networks, while some multi-network products instead issue separate share classes by network; each technical deployment can also require its own token-level identifier. Its terms of ownership may be partly written in legal documents and partly enforced through smart-contract logic that an authorized party can amend. Its authoritative NAV may be struck once daily even when the token can be transferred or traded on-chain outside the primary market's operating hours.

Three themes run through the pages that follow.

- **Where the gap sits.** The settlement technology works. What is missing is the structured, governed, attested data that lets an institution describe what it holds, to whom it has recourse, under what conditions it can exit, and who has the power to interfere. This is a reference data question rather than a blockchain infrastructure question.
- **Extension rather than replacement.** Legacy security masters, accounting books of record, and regulatory reporting stacks are load-bearing and are not going away (at least not in the foreseeable future). A data model that requires abandoning the existing golden record is unlikely to be adopted, regardless of its elegance. The practical work is additive: new identifiers alongside existing keys, extended enumerations within existing fields, and a maintained mapping between the blockchain-native and institutional identifier worlds.
- **The connective tissue.** That connection (the bidirectional, maintained linkage between tokenized contract addresses and the traditional instrument, between networks and venue identifiers, and between the growing cast of tokenization participants and their Legal Entity Identifiers) is the component with the least natural owner. Issuers publish contract addresses. Numbering agencies assign ISINs. Joining the token to the instrument is nobody's job by default.

Leaving the gap open does not produce a visible failure. The security master goes on answering every query put to it, and the answers are simply wrong.

1. The growth of tokenized real-world assets

1.1 Where the volume actually sits

The tokenized RWA market is not evenly distributed, and the concentration matters for data purposes.

The largest share of assets, at least today, sits in **tokenized cash equivalents**: government money market funds, treasury-backed vehicles, and short-duration cash management products issued by established asset managers. These are the instruments most likely to appear on an institutional balance sheet in the near term, precisely because they are the most familiar in legal form.

Adjacent but distinct are stablecoins, which are frequently grouped with tokenized RWAs. For security-master modeling, they should not be treated as the same product class as tokenized fund shares. Their legal claim structure, redemption mechanics, and regulatory characterization differ materially from a fund share class, and treating them as a single category is the first modeling error many firms make.

Beyond that: **private credit and fund interests**, where tokenization is used primarily for transfer efficiency and secondary liquidity in assets that historically had neither; **tokenized equities and equity-like exposure**, where the legal claim varies enormously between products that look superficially identical; **commodities**, predominantly gold, with one of the longer operating histories among tokenized RWA categories; and an emerging set of **collateral, repo, and deposit token** applications where the institutional interest is in intraday mobility rather than in the asset itself.

Tokenized Money Market Fund Growth

Source: Lukka. Excludes treasury or government bond funds that hold longer-duration instruments.

1.2 Why this cycle is structurally different

Three changes distinguish the current wave from earlier tokenization activity, and each one increases rather than decreases the reference data burden.

The issuers are incumbents. The products gaining institutional traction are issued by large asset managers, banks, and market infrastructure providers, not by startups. This is generally read as a de-risking signal. For data purposes it is the opposite of simplifying: incumbent issuance means these instruments enter the book through the front door, subject to the full control framework, rather than through a ring-fenced digital assets carve-out where bespoke handling is tolerated.

Largest 5 Traditional Money-Manager issuers by AUM.

Traditional money-manager issuers by AUM

Manager	Tokenized AUM*
BlackRock	$2.74B
Franklin Templeton	$2.46B
JPMorgan	$897M
Invesco	$821M
WisdomTree	$766M

***AUM shown refers to the tokenized fund products included in this analysis, not each manager's total firmwide AUM. As of August 13, 2026.*

The products are wrappers on regulated instruments. Earlier cycles produced novel synthetic exposures with no analog in existing taxonomies. The current cycle produces something harder: instruments that map almost perfectly onto existing categories. A tokenized money market fund is a money market fund. The temptation is to record it as one and move on, which works until the first contract migration, the first blocked transfer, or the first auditor question about position verification.

Regulatory scaffolding now exists. Several major jurisdictions have established or are implementing frameworks relevant to digital-asset issuance, custody, and market conduct.This enables institutional participation and simultaneously creates a new data requirement: regulatory characterization is now jurisdiction-specific, sourced, and dated, rather than an open question everyone agrees to defer.

Example: Regulatory classifications

Lukka Asset Identifier	Asset Name	Regulatory Classification
LA5BB0D4984	WisdomTree Private Credit and Alternative Income Digital Fund	SEC: Issuer Sponsored
LA4E4B27174	Mirrored Alibaba	SEC: 3rd Party Synthetic
LA05A9E1C27	MARA Holdings (Ondo Tokenized Stock)	SEC: 3rd Party Custodial
LA6EV2NKQ95	Bitcoin	SEC: Digital Commodities
LA8BS7Y9L42	CryptoPunks 9991	SEC: Digital Collectibles
LA1S39UVDK1	USD Coin	SEC: Stablecoins
LA7510GAWX9	Ethereum Name Service 266855850177000196867858450568273546622636240544518818396223339049987785 4058	SEC: Digital Tools

ⓘ ***Note about this data:*** *as part of Lukka's Terms & Conditions dataset, Lukka tracks a variety of fields that help place a digital asset into regulatory guidance and published principles, such as MiCA. For this example we show the "Classification of Crypto Assets" from the SEC's "Application of the Federal Securities Laws to Certain Types of Crypto Assets and Certain Transactions Involving Crypto Assets."*

1.3 The consequence for data infrastructure

The innovation in tokenization happened at the settlement layer. The burden landed on the reference data layer.

A single fund share class can now exist as multiple simultaneous deployments across multiple networks, each with its own operational characteristics, all representing the same legal claim. The same underlying economic exposure can reach an institution through a native token, a wrapped version of that token, and a third-party vehicle issuing notes referencing the fund: three positions in three systems that should aggregate to one exposure and, in most current implementations, do not.

Example: Digital assets that provide economic exposure to gold via different mechanisms

Lukka Asset Identifier	Asset Name	Underlying Asset	Is Collateral	Is Reference	Value Proposition
LA12AGMPCH8	PAX Gold	Gold	✔	✘	Commodity Exposure: Gold
LA4556E0295	Robinhood Tokenized GLD	SPDR Gold Shares	✔	✘	Commodity Exposure: Gold
LA1G97NSPX8	Synth sXAU	Gold	✘	✔	Commodity Exposure: Gold

ⓘ ***Note about this data:*** *as part of Lukka's Terms & Conditions dataset, Lukka tracks the underlying asset, whether that asset serves as collateral or only as a reference asset, and the economic exposure provided by the instrument. This table contains several fields relevant to communicating that concept.*

None of this is a settlement problem. Every one of these instruments settles. The question is whether the institution holding them can describe them.

1.4 The rails are not being replaced

It has become conventional to describe tokenization as disintermediating the traditional post-trade stack. Whatever the long-run merits of that argument, it is not a useful planning assumption, and firms that have built programs around it have generally not shipped anything.

The security master, the accounting book of record, the custodian reconciliation process, and the regulatory reporting stack are staying for any planning horizon that matters. The reasons are not sentimental. They carry decades of embedded control frameworks that have been examined by regulators and signed off by auditors. They are understood by the people who operate them. And tokenized holdings are, for nearly every institution, a rounding error against the total book those systems service. Nobody is going to re-platform a firm's golden record to accommodate a fraction of a percent of assets, nor should they.

More fundamentally: the instruments are new, but the obligations attached to them are not. A tokenized fund holding must be valued under the same standards, disclosed in the same reports, reconciled to the same standard of evidence, and defended to the same auditors as any other holding. Those obligations run through the existing stack.

This produces the corollary that governs the remainder of this paper. Disintermediation was the thesis; integration is the deployment. Any reference data model that requires replacing the existing golden record will not be adopted. The work is to extend it and, on the other side, to require that issuers and networks emit data in a form the existing stack can ingest.

---

2. How security master systems need to change

2.1 The two-sided obligation

It is easy to frame this as a legacy technology problem. That framing is both unfair and unhelpful.

Institutions do carry an obligation: to extend their data models to represent attributes their schemas never anticipated, and to stop treating tokenized instruments as exceptions handled in spreadsheets outside the golden record.

But issuers and networks carry the reciprocal obligation, and it is currently the less discussed one. A great deal of tokenized product data today exists only as prose in an offering memorandum, as a value readable from a smart contract by someone who knows where to look, or as an announcement in a channel no institutional operations function monitors. Products are being designed with terms that no institutional data model can represent, and the resulting slow adoption is then attributed to institutional conservatism.

The failure mode is not legacy architecture. It is the absence of a translation layer between two worlds that each work adequately on their own terms.

Two principles make the extension work tractable.

Separate the legal layer from the technical layer. Legacy security masters model the legal instrument, and they model it well. Tokenized products require modeling the legal instrument *and* its technical instantiations, independently, because the two do not map one-to-one. This distinction is what makes the change additive: the legal layer already has a home in the existing master. The technical layer is the extension beneath it.

Keep the fields orthogonal. The industry's instinct on encountering a new instrument type is to add a flag (`is_tokenized` = Y) or to create a new asset class. Both collapse independent attributes into a single label, and both fail on the first product that does not fit the implied stereotype. Claim type, issuance basis, backing model, redemption mechanism, transfer restrictions, and custody model vary independently of one another. They have to be captured independently.

The subsections that follow walk the areas where existing models run out of fields, and each closes with the pragmatic bridge, the way to represent the new attribute without rebuilding anything.

2.2 Identity, coverage, and grouping

An ISIN, for instance, identifies a security or financial instrument. It does not inherently identify each blockchain deployment of that instrument.

There is a useful precedent for the shape of the problem. A cross-listed equity has one instrument record with several listing records beneath it, each carrying venue-specific attributes: local ticker, trading currency, settlement convention, lot size. Tokenized deployments fit that pattern. A fund share class deployed across five networks is one legal instrument with five instantiations, which is a familiar enough shape that most institutions assume their existing structures will absorb it.

Where the precedent stops being reassuring is in *what* varies between instances. A listing venue does not change the holder's legal counterparty, does not determine who is eligible to hold the security, does not affect whether the holding is interchangeable with the same instrument elsewhere, and cannot immobilize a position. A network deployment can do all four. Deployments differ in minimum investment, transfer mechanics, eligible counterparties, available liquidity, and whether a holding on one network can be moved to another at all.

Example: Data that is dependent on blockchain deployment

Lukka Asset Identifier	Blockchain	Fee	Fee Type
LA0NAK0BZO4	Ethereum	50 Bps	Management Fee
LA94A401238	Polygon	20 Bps	Management Fee
LA87826BB44	Optimistic Ethereum	50 Bps	Management Fee
LA6C38BAF06	Avalanche Contract Chain (C-Chain)	20 Bps	Management Fee
LA05EA3A704	Solana	20 Bps	Management Fee
LA2447D5F81	Arbitrum	50 Bps	Management Fee
LA851E68316	Aptos	20 Bps	Management Fee
LA0E9D696B1	BNB Smart Chain	18 Bps	Management Fee

ⓘ ***Note about this data:*** *as part of Lukka's Terms & Conditions dataset, data is modeled at the asset-blockchain level. This table contains several fields relevant to communicating that concept as it relates to Management Fees for the BlackRock USD Institutional Digital Liquidity Fund (BUIDL).*

The identifier problem compounds with derivation. A token may wrap another token, which wraps a fund share. Each layer introduces a distinct issuer, a distinct legal claim, and a distinct set of failure modes, while the ultimate underlying exposure is unchanged. Cross-chain representation introduces the same distinction: an issuer-native deployment may represent the claim directly on another network, while a bridged token may represent a claim on tokens held through bridge infrastructure. The relationship between deployments therefore matters alongside their grouping. Nothing in a conventional master expresses how many layers sit between the holder and the asset.

Example: Digital assets that hold other digital assets that hold other assets

Lukka Asset Identifier	Asset Name	Underlying Asset ID & Name	Further Underlying Exposure
LA4SJGDYZQ4	Ondo Short-Term U.S. Government Bond Fund	LA4830417 37: State Street Galaxy OnChain Liquidity Sweep Fund	Traditional Market Exposure: Money market
		LA0NAK0BZO4:BlackRock USD Institutional Digital Liquidity Fund	Traditional Market Exposure: Money market
		LA0D21DBB21: Fidelity Treasury Digital Fund	Traditional Market Exposure: Money market
		LA756A755B7 Franklin OnChain U.S. Government Money Fund	Traditional Market Exposure: Money market
		LA1S39UVDK1: USD Coin (Ethereum)	USD: U.S.Dollar

Note about this data: *as part of Lukka's Terms & Conditions dataset, Lukka tracks a variety of fields that help customers track the underlying asset(s) of a digital asset. This has applications to Wrapped Assets, Stablecoins, Tokenized RWAs, etc. For this example, we use a recursive process to identify the layers involved so customers can appropriately understand, assess, and make decisions about risk.*

There are gaps in the other direction too. Many tokens lack a generally accepted identifier with sufficiently broad institutional adoption and coverage." Legal Entity Identifier coverage on issuing vehicles and special purpose entities is inconsistent. Existing instrument classification codes were not designed to capture every technical attribute introduced by tokenized instruments

The more consequential gap, though, is between an identifier scheme existing and an identifier scheme covering the market. The population that has to be valued, reconciled, and reported is not determined by what identifiers exist; it is determined by what a bank's clients, a fund administrator's portfolios, and a custodian's accounts actually contain. Coverage is the paramount test here, and it is the one most easily obscured by a compliance-sounding label.

The pragmatic bridge: the key has to be the identifier with complete coverage, because an identifier that is absent for part of the population is not functioning as a key at all. A token with no ISIN still has to be held, valued, reconciled, and reported. A model that cannot identify it until a numbering agency acts has already broken operationally, and the break shows up as a position nobody can name rather than as a missing field.

Standard identifiers belong in the data model, but as attributes on the record rather than as its spine. ISINs and LEIs should be populated wherever they exist, because plenty of downstream uses require them: regulatory reporting, client statements, settlement messaging, and every system already keyed on them. Nothing is surrendered by carrying them this way. What changes is that their absence no longer prevents an instrument from being represented at all. Coverage is the first requirement of an identifier, and pedigree is the second.

None of this asks an institution to re-key its own master. Existing keys stay exactly where they are. It is a statement about what the reference data layer must be keyed on in order to deliver reliably into a master that keeps its own.

Rather than subordinating deployments to a parent record, give each one its own identifier. A deployment on a given network at a given contract address is a distinct thing to hold, and it should be identifiable in its own right. This is also what lets an instrument be recorded the moment it is observed, before anyone has determined which fund it belongs to or whether an ISIN exists for it, which a parent-child structure cannot do because a child requires a parent to hang from.

Sameness is then expressed by a second identifier that groups deployments representing the same underlying instrument across networks. One fund share class deployed on five networks yields five deployment identifiers and one group identifier. Risk aggregates on the group, settlement instructions resolve to the deployment, and neither has to understand the other's level.

Example: Accounting for Cross Chain Asset Deployments

Lukka Asset Identifier	Asset Name	Blockchain	Asset Group ID / Name
LA0NAK0BZO4	BlackRock USD Institutional Digital Liquidity Fund	Ethereum	10213 / BUIDL
LA94A401238	BlackRock USD Institutional Digital Liquidity Fund	Polygon	10213 / BUIDL
LA87826BB44	BlackRock USD Institutional Digital Liquidity Fund	Optimistic Ethereum	10213 / BUIDL
LA6C38BAF06	BlackRock USD Institutional Digital Liquidity Fund	Avalanche Contract Chain (C-Chain)	10213 / BUIDL
LA05EA3A704	BlackRock USD Institutional Digital Liquidity Fund	Solana	10213 / BUIDL
LA2447D5F81	BlackRock USD Institutional Digital Liquidity Fund	Arbitrum	10213 / BUIDL
LA851E68316	BlackRock USD Institutional Digital Liquidity Fund	Aptos	10213 / BUIDL
LA0E9D696B1	BlackRock USD Institutional Digital Liquidity Fund	BNB Smart Chain	10213 / BUIDL

ⓘ ***Note about this data:*** *as part of Lukka's Terms & Conditions dataset, we assign unique identifiers at the asset-blockchain level. This table contains several fields relevant to communicating that concept in the data model.*

The trade-off is worth stating plainly. A hierarchy makes aggregation implicit, whereas a group identifier makes it explicit, and a consuming system that neglects the join will see five unrelated positions rather than one exposure.

2.3 Venue, settlement, and the missing depository

Market Identifier Codes describe trading venues, and they carry descriptive attributes alongside the code: operating entity, country, venue type, status. Identifier schemes for tokens and networks have been defined, including under an ISO standard. What has not arrived is either broad adoption or a comparable attribute set, and institutions consequently find that the network, functionally the venue where a position lives and moves, is the thing their venue tables have least to say about.

Example: Standardized "location" identification embedded into the reference data record

Lukka Blockchain Identifier	Blockchain	Lukka Asset Identifier	Asset Name
LBCCQ72TLA0	Aptos	LA001F111B8	Franklin OnChain U.S. Government Money Fund (BENJI)
LB230K0C5V4	Arbitrum	LA644EE8875	Franklin OnChain U.S. Government Money Fund (BENJI)
LB0ASLG5870	Avalanche Contract Chain (C-Chain)	LA54B873F07	Franklin OnChain U.S. Government Money Fund (BENJI)
LB1E6F055B4	BASE Chain	LA360E9E680	Franklin OnChain U.S. Government Money Fund (BENJI)
LB5BNEJF816	Ethereum	LA756A755B7	Franklin OnChain U.S. Government Money Fund (BENJI)
LB325JNLUL2	Polygon	LA9A45P9AI1	Franklin OnChain U.S. Government Money Fund (BENJI)
LB3MLKKRGX5	Solana	LA2D8C76E61	Franklin OnChain U.S. Government Money Fund (BENJI)
LB3VCFP5HK1	Stellar	LA9I6OV25L9	Franklin OnChain U.S. Government Money Fund (BENJI)

ⓘ ***Note about this data:*** *as part of Lukka's Terms & Conditions dataset, we assign unique identifiers for a blockchain. This table contains several fields relevant to communicating a standardized blockchain identification. In this case it is the Franklin OnChain U.S. Government Money Fund (BENJI). Note: iBENJI, sgBENJI, and grBENJI are excluded here.*

This is the gap the industry has been slowest to recognize, because the existence of a code is easily mistaken for coverage of the market. A firm checking whether network identification is solved will find that it is, in the sense that a scheme exists. It discovers the difference between a scheme existing and its own holdings being in it considerably later, usually at the point where a report has to be produced, a position has to be grouped by venue for risk, or a settlement policy has to name the network it applies to. By then the workaround is in place and the exception process is staffed.

Settlement conventions break more sharply. Legacy settlement fields often express a settlement cycle in a number of days. Atomic settlement does not fit that convention. Atomic settlement is not a number of days. Neither is probabilistic finality, where a transfer becomes progressively rather than instantaneously final, and where the institution's own policy on how many confirmations constitute settlement is a data-driven decision with no field to hold the inputs.

Position location changes character entirely. There is no account at a central depository. There is a wallet address, and the evidentiary chain that proves the address is yours.

The pragmatic bridge: treat networks as reference data records in their own right (created once, referenced everywhere), exactly as exchanges and depositories are treated today. Carry a standards-based network identifier on that record wherever one has been assigned, and do not let its absence stop the network from being represented, which is the same discipline section 2.2 applies at the instrument level. For settlement, extend the existing settlement model to capture both settlement mechanics and finality characteristics rather than forcing blockchain settlement into a day-count convention.

2.4 Legal claim structure

This is the area where superficially identical products diverge most severely, and where the cost of getting it wrong is highest.

A holder may have direct legal title to the underlying asset; a beneficial interest held through a nominee or transfer agent; a contractual claim against an issuing vehicle that itself holds the asset; or purely synthetic exposure with no claim on anything. All four can be presented in marketing materials as a tokenized version of the same fund.

Example: Holder claim types

Lukka Asset Identifier	Asset Name	Holder Claim Type
LA12AGMPCH8	PAX Gold	Direct Ownership
LA4B67C11C3	Apollo Diversified Credit Securitize Fund	Indirect Beneficial - SPV Wrapper
LA26AB103F7	Dinari AMD	Indirect Beneficial - Custodial Depositary
LA3A20D7A71	Amazon xStock	Contractual - Issuer Debt Claim
LA1G97NSPX8	Synth sXAU	Contractual - Pure Synthetic / Swap

ⓘ ***Note about this data:*** *as part of Lukka's Terms & Conditions dataset, Lukka tracks and classifies the holder claim type to the underlying asset(s) so users can understand risk and power pre-trade compliance and post-trade risk management use cases. This table contains several fields relevant to communicating that concept in the data model.*

Economic entitlement sits alongside the claim itself. Two tokens over the same pool of assets can differ on whether the holder receives the yield those assets generate or whether it is retained by the issuer or another party. That is a question about the claim rather than about income mechanics, and it is invisible from the price.

Underneath that sits the question of issuance basis: whether the token was natively issued on-chain as the fund's own share class, or whether it is a depositary-receipt-style instrument over an off-chain register, or a note issued by a third-party vehicle. And underneath that is a question most firms have never had to ask, because the answer was previously always the same: is the on-chain record the legal register of holders, or a mirror of a register maintained elsewhere? If it is a mirror, the chain is not the source of truth, and the direction of reconciliation reverses.

Example: Official register types

Lukka Asset Identifier	Asset Name	Register Location
LA325A37706	4-week T-Bill - DB Series-8 (TBL-20241112-8)	On-Chain
LA9925A3261	JPMorgan Chase (Ondo Tokenized)	Off-Chain
LA96E0813F6	Superstate Crypto Carry Fund	Hybrid

ⓘ ***Note about this data:*** *as part of Lukka's Terms & Conditions dataset, Lukka tracks and classifies the register location to the underlying asset(s) so users can understand location and operational processes for ownership. This table contains several fields relevant to communicating that concept in the data model.*

Jurisdiction splits as well. The issuing vehicle may sit in one jurisdiction, the underlying assets in another, and the token terms may be governed by the law of a third.

Example: LA325A37706 4-week T-Bill - DB Series-8 (TBL-20241112-8)

Issuer Jurisdiction	Security Agent Jurisdiction	Governing Law	Underlying Issuance
Bermuda	New Hampshire (USA)	English Law	United States of America

ⓘ ***Note about this data:*** *as part of Lukka's Terms & Conditions dataset, Lukka tracks and classifies the various participants involved in a digital asset. Those participant legal names are stored along with the legal entity type and incorporation jurisdiction.*

The pragmatic bridge: most of these are attributes of the legal instrument and belong where the master already holds issuer, guarantor, governing law, and country of risk. The precedent is closer than it looks. Depositary receipts, participation notes, and exchange-traded notes are all claim-structure variants that existing masters distinguish today, and jurisdiction of the vehicle against jurisdiction of the underlying assets is already a standard pair of fields. What is new is that claim structure can no longer be inferred from instrument type, so it has to become an explicit field instead of an implication of the type code.

Two of these attributes do not belong in the legal tier at all. Whether the on-chain record is the authoritative register, and the basis on which the instrument was issued, are both statements about how the legal instrument relates to a particular technical instantiation, and both can vary across deployments of the same share class. These belong on the deployment record introduced in section 2.2, or more precisely at the join between the two layers. An attribute that describes how the layers relate cannot be stored inside either one.

2.5 Backing and reserves

Where an instrument is not a fund share class, the question of what stands behind it becomes a first-class data problem: backing model, reserve composition, the eligible asset policy governing that composition, the entity actually holding the reserves, and the jurisdiction it holds them in.

Example: Backing Model Classifications

Lukka Asset Identifier	Asset Name	Backing Model
LA12AGMPCH8	PAX Gold	1:1 Asset-Backed
LA1LC6NTDP7	Multi-Collateral Dai	Over-Collateralized
LA6H5DBYF65	TerraUSD	No Backing
LA6EV2NKQ95	Bitcoin	Not Applicable

ⓘ ***Note about this data:*** *as part of Lukka's Terms & Conditions dataset, Lukka tracks and classifies the backing model so users can understand risk. This table contains several fields relevant to communicating that concept in the data model.*

The attestation layer deserves more attention than it usually gets. A monthly agreed-upon-procedures report and a continuously verifiable proof of reserves are both described as "attested." They are not the same control, and the difference is not visible unless the attestation type, provider, cadence, and reporting lag are all captured. A reserve report that is current as of six weeks ago is a materially different input to a credit decision than one current as of yesterday.

Yield mechanics deserve equal attention. Economic return may reach the holder through a rebasing token supply, through accrual in NAV, through a separate distribution token, or not at all, because the yield is retained by the issuer. Two tokens over identical assets can differ entirely on this dimension, and the difference is invisible from the price.

Example: Accounting for Distribution

Lukka Asset Identifier	Asset Name	Use of Income	Frequency
LA167C21277	WisdomTree TIPS Digital Fund	Cash Distribution	Monthly
LA6C38BAF06	BlackRock USD Institutional Digital Liquidity Fund	Rebase Distribution	Daily

ⓘ ***Note about this data:*** *as part of Lukka's Terms & Conditions dataset, we track various fields around the distribution of income. This table contains several fields relevant to communicating that concept in the data model.*

The pragmatic bridge: the static attributes extend structures the master already holds. The backing model, the eligible-asset policy, the reserve custodian, and the attestation type, provider and cadence sit alongside the collateral and guarantee information already carried for asset-backed and covered instruments. Reserve

composition itself is not a static field and should not be forced into one. It is periodic disclosure, handled the way servicer reports and cover pool data are handled today, with the master holding the reference and the as-of date rather than the contents.

Yield needs to be split across three places instead of one. Whether return accrues in value or is distributed maps directly onto the accumulating and distributing share class treatment that already exists. Whether the holder is entitled to the yield at all is a claim question rather than an income question, and belongs with the economic entitlement attribute in section 2.4. And a rebasing supply is neither of those: it is a change in unit quantity, which makes it a lifecycle event under section 2.9. Recording it as an income attribute is how a position's unit count changes without the accounting engine ever being told.

2.6 Redemption and liquidity mechanics

Legacy liquidity fields assume that liquidity is a property of trading: volume, spread, market depth. For tokenized RWAs, liquidity is substantially a property of contractual mechanics.

Whether a holder can redeem at all is the first question, and the answer is frequently no: redemption rights may sit only with authorized participants, meaning ordinary holders exit through someone else's balance sheet. Beyond that: which channels exist, what asset redemption proceeds arrive in, what notice period and cutoff apply, what gates or queues can be imposed, and what minimums govern. Minimums differ by network deployment for the same share class, so a single instrument-level minimum is wrong for most multi-network products.

Example: Redemption

Lukka Asset Identifier	Asset Name	Redemption Mechanism
LA6C38BAF06	BlackRock USD Institutional Digital Liquidity Fund	Direct Issuer Redemption
LA167C21277	WisdomTree TIPS Digital Fund	Direct Issuer Redemption
LA536006430	Backed IB01 $ Treasury Bond 0-1yr	Direct Issuer Redemption
LA1G97NSPX8	Synth sXAU	No Redemption

ⓘ ***Note about this data:*** *as part of Lukka's Terms & Conditions dataset, we track data around the redemption mechanisms. This table contains data relevant to communicating that concept in the data model.*

Then there is the structural mismatch that defines this asset class. The token trades continuously. The primary market operates during business hours, on a business day calendar, in a specific timezone. An official NAV, a subscription or redemption price, and an observable secondary-market price can therefore coexist without representing the same thing. The master needs to preserve that price-source hierarchy rather than collapse them into competing observations of a single field. That mismatch is the mechanical origin of persistent premium and discount behavior.

One further dependency belongs here: if redemption settles in a stablecoin rather than fiat, that stablecoin's issuer is a counterparty in the redemption path and belongs in the entity graph.

Example: Redemptions, subscriptions and distributions

Lukka Asset Identifier	Asset Name	Mechanism	Subscription	Redemption	Distribution
LA10A41FDA8	1-month T-Bill - DB Series-3 (TBL-20240618-3)	Fiat	✘	✘	✘
		Crypto	✔	✔	✔

ⓘ ***Note about this data:*** *as part of Lukka's Terms & Conditions dataset, we track data around the currency types supported during redemptions, subscriptions and distributions. This table contains data relevant to communicating that concept in the data model.*

The pragmatic bridge: most of this extends the terms-and-conditions structures that already exist for open-ended funds and structured notes. The new requirement is granularity: several of these values must be recorded per deployment rather than per instrument.

2.7 Terms encoded in code

This is the area with the least precedent and, arguably, the greatest consequence.

In a tokenized instrument, some of the terms of ownership are not in a document. They are in a smart contract, enforced automatically, and amendable by a party the terms designate. Transfer eligibility may be enforced at the protocol level against an allowlist. And that designated party may hold the ability to pause transfers, freeze a specific address, force a transfer or claw back tokens, burn from a holder's balance, mint additional supply, change the price oracle, or upgrade the contract logic outright.

Example: Transfer eligibility

Lukka Asset Identifier	Asset Name	Transfer Restrictions	Transfer Gate Location
LA6C38BAF06	BlackRock USD Institutional Digital Liquidity Fund	Permissioned Whitelist	On-Chain at Every Transfer
LA167C21277	WisdomTree TIPS Digital Fund	Permissioned Whitelist	On-Chain at Every Transfer
LA536006430	Backed IB01 $ Treasury Bond 0-1yr	Freely Transferable	At Issuance & Redemption
LA20C12E397	Toucan Protocol: Nature Carbon Tonne	Freely Transferable	At Redemption Only

ⓘ ***Note about this data:*** *as part of Lukka's Terms & Conditions dataset, we track data around transfer eligibility. This table contains several fields relevant to communicating that concept in the data model.*

Each of those is a discrete power exercisable over an institution's position. The legal and technical layers can also diverge: documentation may describe a power more narrowly than the deployed contract technically permits, or the reverse. The master

therefore needs to preserve legally stated rights and technically observable capabilities separately. Whether the contract is upgradeable, whether upgrades are subject to a timelock, and how long that timelock runs together determine whether a holder has any window to react at all. A legacy terms-and-conditions table may have no field to record that an authorized administrator can unilaterally immobilize a holder's position, effective immediately.

Two related attributes are worth capturing and almost never are. First, whether these powers have ever actually been exercised, and when. Capability and demonstrated willingness to use it are different risks, and only one of them appears anywhere in current data models. Second, whether the token may be pledged at all under its terms. Some prohibit it outright. Collateral desks assume permission by default, because no field says otherwise.

There is also an operational attribute hiding here: how long it takes to get a new wallet approved to receive a permissioned token. That is a liquidity input. Treating it as an operations detail rather than a data field is how settlement dates get missed.

The pragmatic bridge: model these as an enumerated capability set attached to the deployment record, in the same way negative covenants and issuer call provisions are already modeled for debt instruments. The concept, namely enumerated issuer rights over a holder's position, is not new to the security master. Only the specific rights are.

2.8 Custody and position location

Custody arrangements for tokenized assets span qualified custodians, direct institutional custody, and shared-control models with no clean legacy analog. The reconciliation target shifts from a custodian statement to the state of a network.

Two attributes matter more than they appear to. Address attribution, meaning how an institution demonstrates that a given address is in fact its own, is the first question an auditor asks and has no legacy home whatsoever. And whether a position can be independently verified from network state without the issuer's cooperation is a meaningful control strength that distinguishes products which otherwise look identical on paper.

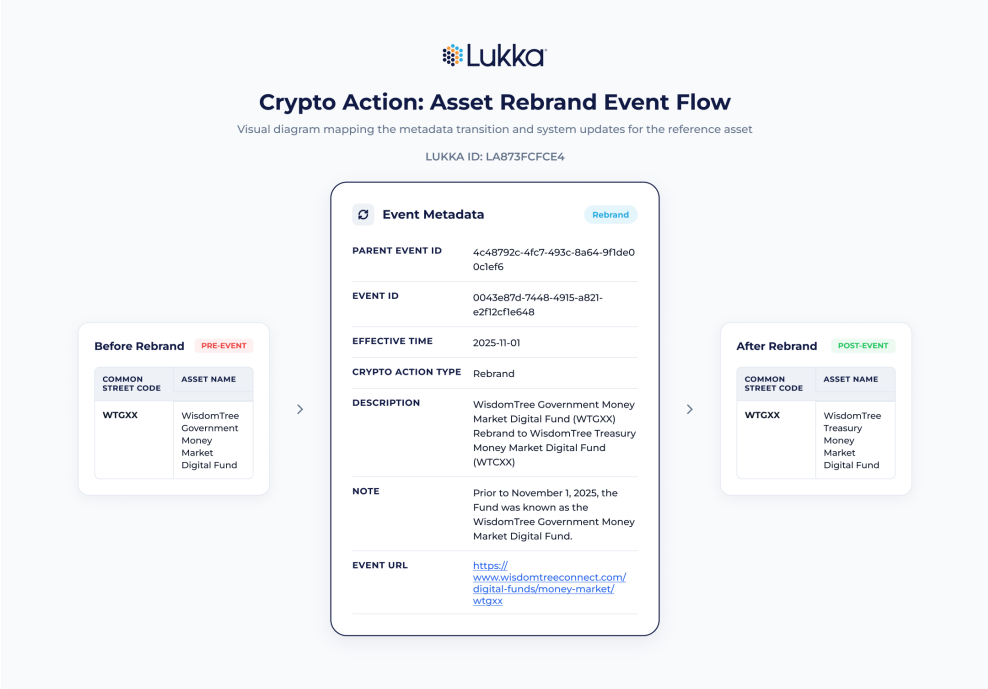
A final dependency: moving or redeeming a position may require holding a separate network asset to pay transaction fees. A lack of the required network asset can delay or prevent a transfer until transaction-fee requirements are met. Nothing in a security master anticipates it.

The pragmatic bridge: network state becomes an input to the existing custodian reconciliation process rather than the basis of a parallel process. One reconciliation, one break report, one set of controls, with an additional evidence source.

2.9 Lifecycle events and corporate actions

Tokenized instruments experience events that legacy corporate action taxonomies have no code for: contract migration to a new address, redeployment, chain deprecation, bridge deprecation, network forks, token standard upgrades, supply rebasing and other changes to unit count, redenomination, eligibility policy changes, and administrative actions actually executed against holders.

A contract migration is, economically, a mandatory exchange. There is simply no announcing agent. The notification may be a press release, an on-chain event, a documented feed, or a post on a social platform. Recording *which* source an institution is relying on is the difference between a control and a habit. The gap between notice given and notice contractually required is, incidentally, a measurable issuer quality metric, and one that nobody currently measures.



The pragmatic bridge: the event types are new; the distribution mechanics do not have to be. Map each new event type to the closest existing corporate action category and deliver it through existing channels, so a contract migration reaches the accounting engine as a mandatory exchange rather than as an email that someone forwards.

2.10 Regulatory characterization and classification

Two classifications are required, recorded independently: the legal form and the technical form. This is the dual-layer principle made concrete, and collapsing the two

(which is what a single "digital asset" asset class does) destroys information that is needed downstream.

Regulatory characterization is jurisdiction-specific. The same token may receive materially different regulatory classifications across jurisdictions. For example, treatment as a security or another regulated financial or payment instrument, which means characterization must carry a source and an as-of date rather than being asserted once.

The classification question institutional buyers most want answered, and which currently has no structured answer anywhere, is eligibility for regulated uses: whether the instrument qualifies for high-quality liquid asset treatment, whether it is accepted as margin or collateral at a clearing venue, whether it is an admitted asset for insurance purposes, and how it is treated for bank capital. These determine whether a product is usable at all by a regulated institution.

The pragmatic bridge: legal form classification uses the vocabulary already in the master. On the technical side, resist the instinct to build one large classification. Most technical attributes are independent controlled lists rather than positions in a tree: token standard, permissioning model, canonical or derived instance, deployment status. Modeling them as separate enumerations is both simpler and more faithful to how they actually vary, and there is precedent in the master already, since CFI encodes each of its positions as its own controlled list. Reserve hierarchy for the one place where it earns its cost, which is classifying what the instrument is, where consumers need to aggregate at a level above the leaf. The test is whether anything downstream needs to roll up. Where it does, a taxonomy. Where it does not, an enumeration, and the discipline to extend it without breaking the consumers who read it.

2.11 Participants and the entity graph

Most commentary on tokenization reference data stops at instrument identifiers. The participant layer is the larger gap.

A legacy security master models a small number of related parties: issuer, guarantor, perhaps a paying agent. A tokenized fund has an issuer, a sponsor, an investment manager, a transfer agent, a token administrator, a custodian of the underlying assets and possibly sub-custodians, a paying and redemption agent, a distributor, an allowlist administrator, an oracle provider, a financial statement auditor, a reserve attestor, a smart contract auditor, potentially a bridge operator, and a network operator. Frequently there is also a special purpose vehicle that is the holder's actual legal counterparty.

Each of those needs an entity record, a Legal Entity Identifier or an explicit reason one does not exist, a jurisdiction, and effective dates. Without that, counterparty exposure aggregation is guesswork, and it will usually aggregate to the brand name on the product rather than to the entity the claim actually runs against.

Example: Linking the LEI to the tokenized asset participant list

Participant Type	Participant Name	LEI
Auditor	PRICEWATERHOUSECOOPERS LLP	5493002GVO7EO8RNNS37
Custodian	DB International Trust (Singapore) Limited	529900PZS6JN7D8F1A13
Tokenization	Franklin Templeton Investor Services, LLC	(no LEI Issued)
Primary Issuer	FRANKLIN TEMPLETON INVESTMENTS VCC	2549006B0QZEBXY9Q594
Advisor	Franklin Resources, Inc.	RIFQSET379FOGTEFKS80
Fund Administrator	Deutsche Bank AG, Singapore Branch	(no LEI Issued)
Investment Manager	TEMPLETON ASSET MANAGEMENT LTD	7AYVD3NQFQ1OTZURH567
Distributor	Franklin Distributors, LLC	(no LEI Issued)
Sub-Advisor	Western Asset Management Company, LLC	549300C5A561UXUICN46

ⓘ ***Note about this data:*** *as part of Lukka's Terms & Conditions dataset, we track participants involved in digital assets. Lukka also ingests and processes the LEI file to enrich its entity mapping dataset. This table contains several fields relevant to communicating the mapping of the entity back to its Legal entity Identifier. In this case it is the Franklin OnChain U.S. Dollar Short-Term Money Market Fund (sgBENJI).*

Two attributes are worth singling out. **Role concentration** (where one entity holds multiple roles that convention would separate, such as issuer and token administrator) is a control finding waiting to happen, and there is no field for it anywhere today. And **role substitutability** (whether a role can be replaced without holder consent or notice) determines whether a counterparty record has a shelf life.

The pragmatic bridge: the entity master already exists and already carries LEIs. What is required is an expanded role taxonomy and the discipline of populating it. This is the participant half of the connective tissue, and it is where the effort pays the highest return in risk aggregation quality.

3. The risks of not tracking these fields properly

The risks below are grouped by where the consequence lands. Each is short, because the point is the consequence rather than the mechanism.

3.1 Risk aggregation blindness

The same underlying exposure held through three wrappers across four networks aggregates to nothing if the data model has no way to relate them. Concentration becomes invisible in exactly the dimensions that matter for tokenized assets: single issuer, single network, single bridge, single stablecoin leg in the redemption path. Counterparty exposure maps to the recognizable asset manager when the

enforceable claim actually sits against a special purpose vehicle. Liquidity risk is misstated in both directions, because redemption mechanics are prose in a document rather than structured data in a field.

3.2 Compliance and regulatory exposure

Where transfer eligibility is enforced in code and absent from the pre-trade check, the institution discovers the restriction at the point of a failed or blocked transfer, or, worse, completes a transfer to an ineligible party and creates a breach. Sanctions and financial crime screening now has an address-level dimension that must be linked to the instrument, participant, and account data maintained elsewhere in the institutional stack. Jurisdictional characterization errors drive incorrect regulatory treatment and reporting, and tax and information reporting obligations turn on legal form attributes that were never captured.

3.3 Collateral and financing

An institution can pledge a permissioned token that its lender is unable, legally or technically, to seize on default. Haircuts and eligibility criteria get set without reference to redemption channel, transfer restriction, or the possibility that pledging is prohibited outright by the instrument's terms. These enforceability assumptions hold until they are tested, which is by definition the worst moment to discover the answer.

3.4 Operational and settlement

Transfers are irreversible and can be directed to the wrong address or the wrong network. Contract migrations get missed, leaving positions in deprecated tokens that may no longer be redeemable. Reconciliation breaks accumulate, and operations teams resolve them in spreadsheets outside the golden record, which is how a temporary workaround becomes permanent infrastructure.

3.5 The compounding cost

Each of these produces a workaround, and workarounds become permanent exception processes with their own control requirements and their own staff. More importantly, retrofitting a golden record after positions already exist is materially harder than designing for them beforehand: the migration has to reconcile history that was recorded under the wrong model.

Tokenized assets do not break the security master, and that is precisely the risk. The master keeps returning answers to every query, on time and in the expected format. They are simply wrong, and nothing in the system indicates it.

3.6 The reality of the current state

There is a second set of risks, less discussed because they arise from how firms have already responded rather than from neglect.

The most common response has been **parallel infrastructure**: a separate book of record, or something close to one, standing up alongside the main stack to hold what the main stack could not represent. This is easy to criticize and worth being honest about: it is where many institutions and market participants actually sit today, and for defensible reasons. The instruments arrived before the data model did. Waiting for a schema change cycle was not an option when a client wanted an allocation this quarter. Firms built what they needed.

The cost is real. Two environments mean permanent reconciliation expense, two versions of the position, and the possibility that the wrong one is cited to a regulator or an auditor. But the useful conclusion is not that these firms should tear down what they built. It is that **the parallel environment has to be joined back to the golden record, and the join is exactly the connective tissue described in section 4.2.** Without a maintained mapping between contract addresses and instrument identifiers, and between tokenization participants and the entities already in the entity master, the two environments cannot be reconciled at all, which is what converts a temporary bridge into a permanent second set of books.

The same logic that says legacy systems are not going away applies here in reverse. The digital asset environment a firm has already built is not going away either. Connecting the two is a tractable program. Demolishing either one is not.

Three further failure modes are worth naming. **Rip-and-replace programs** that stall in procurement and architecture review while positions continue to accumulate in spreadsheets. **Product design that outruns representability**, where issuers create terms no institutional data model can express and then attribute slow adoption to conservatism. And **standards fragmentation**, where each firm invents private conventions for the same fields, ensuring the mapping problem is solved separately by everyone and consistently by no one.

4. What the target state looks like

4.1 Design principles

The target state is describable independently of any provider, and it is worth stating that way first.

Additive schema design. New records and identifiers alongside the keys a firm already uses, and extended enumerations within existing fields, never parallel instrument masters. The objective should be to minimize downstream change to the trial balance, accounting engine, and regulatory reporting stack by supplying richer inputs through existing interfaces wherever possible.

Flat identity, explicit grouping. Each deployment carries its own identifier and is recorded independently, with a separate group identifier expressing which deployments represent the same underlying instrument across networks. Aggregation happens by joining on the group rather than by walking a tree, which keeps a newly observed contract representable before its parentage has been established. Consuming systems resolve to whichever level they need and are not required to understand the other.

Orthogonal fields. Claim type, issuance basis, backing model, redemption mechanism, transfer restriction, and custody model captured independently, because they vary independently.

Dual-layer classification. Legal form and technical form recorded separately. Legal form uses the vocabulary the master already has; technical form is mostly a set of orthogonal enumerations, with hierarchy reserved for the classification consumers actually need to aggregate on. Neither is a flag.

Coverage decides the key; standards ride along. LEI, ISIN, CUSIP, CFI, FIGI, and established messaging standards should be carried wherever they exist, because a great many downstream uses require them. But they cannot be what the model depends on to identify an instrument, because their coverage of this population is incomplete and will remain so for some time. Treat adherence to a standard as a statement about format and governance rather than about coverage: a scheme can be correctly specified, properly sponsored, and still fail to contain the instruments a firm actually holds. Extend standards where they reach, carry them as attributes where they do not, and invent only what nothing else supplies.

Maintained and evidenced. Material attributes and mappings should carry effective dates and source provenance, preserving not only what is true now but what was true at the relevant point in time and the evidence on which that determination was based.

4.2 The connective tissue

Blockchain-native fields on their own produce a well-described island. What makes them usable is the linkage back to the institutional identifier world, which is the connective tissue between the two.

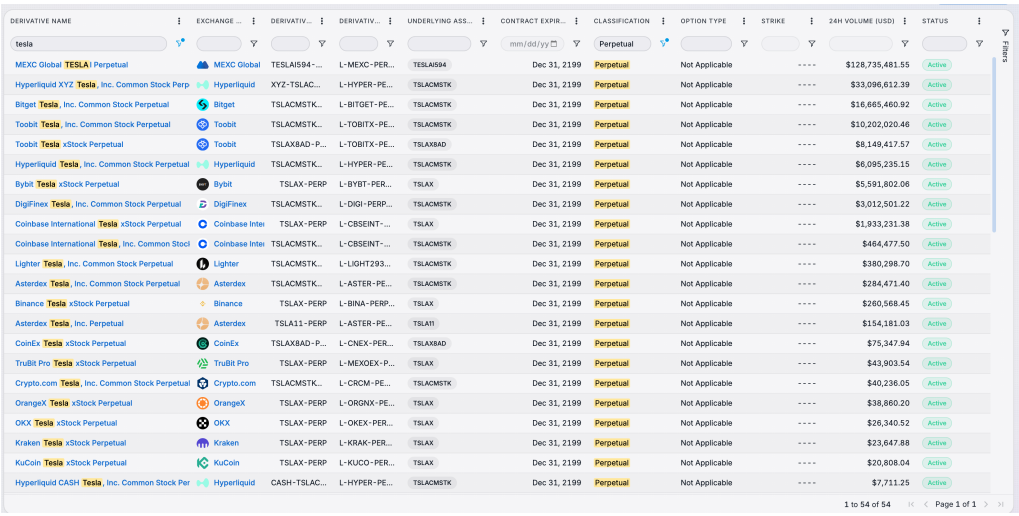
There are three distinct mappings involved, and they are best understood separately.

Instrument mapping. The deployment identifier on one side, the traditional instrument on the other, with the reference underlying asset as the join between them. Many deployments resolve to one group, and the group is what carries the ISIN, CUSIP, FIGI, and CFI wherever they exist. Maintained bidirectionally and at deployment granularity rather than collapsed to the instrument, because a mapping that discards which deployment is actually held has discarded the information that made it necessary.

Participant mapping. Every role in the tokenization structure (issuer, sponsor, manager, transfer agent, token administrator, custodian, redemption agent, allowlist administrator, oracle provider, attestor, bridge operator) carrying an LEI, including the roles for which legacy masters have no slot. This is where counterparty aggregation either works or does not, and it is the least contested gap in the market today.

Network and venue mapping.. Blockchain networks and on-chain trading venues each carry standardized identifiers appropriate to their function and are mapped into the institutional location and venue model without conflating the two.

The layer also carries attributes, not only identifiers. Some facts about a tokenized instrument, such as whether the on-chain record is the authoritative register, describe how the legal instrument relates to a particular technical instantiation. They belong to neither layer on its own and are properties of the join.



Alongside these sits the event bridge. **Crypto actions**, the tokenization-native lifecycle events described in section 2.9, mapped into conventional corporate action categories and delivered through existing channels, so that a contract migration or chain

deprecation reaches the accounting engine as a structured event rather than as a discovery.

The reason this layer requires a maintainer rather than a standard is that neither side produces it. Issuers publish contract addresses because they must. Numbering agencies assign ISINs because that is their function. The join between them, kept current as deployments are added, contracts are migrated, and roles are substituted, is nobody's default responsibility, and it decays continuously if unowned.

It is also what makes the current state described in section 3.6 workable rather than permanent. Where a firm has already stood up a separate environment for digital assets, the mapping layer is the mechanism by which that environment reconciles to the golden record instead of diverging from it. The objective is not one system. It is one answer, from whichever system is asked.

---

5. Evaluating the options

For firms assessing whether to build this layer or source it, the criteria worth applying are:

Coverage, tested against your own holdings. The only meaningful measure is the proportion of what you hold, or expect to hold, that the provider covers today, assessed against your own position file and your clients', not against the provider's total record count. A scheme can be well-governed, standards-endorsed, and largely empty. Ask for a match rate against real positions before anything else.

Demonstrated pace. New networks, deployments, wrappers, and issuers appear continuously, and historical coverage says nothing about whether a provider will keep up. Ask how long it took to onboard the last several significant products, what the process is for a network that did not exist a year ago, and what the observed lag has been between a contract migration and the record being updated. Track record on pace is more predictive than current field count.

Standards carried, not substituted. Does the provider populate LEI, ISIN, CUSIP, CFI, and FIGI wherever they exist and deliver them alongside its own identifiers, so that nothing downstream has to go looking for them? A provider whose identifiers replace the standard ones creates lock-in. A provider that only supplies the standard ones leaves the uncovered population unrepresented. The test is whether both are present on the same record.

Orthogonal fields. Are independent attributes captured independently, or collapsed into asset-class labels that will fail on the next product?

Participant-level identity. Are all tokenization roles modeled with LEIs, or does coverage stop at the issuer?

Delivered into existing pipes. Does the data arrive in a form the existing security master, accounting book, and risk system can consume, or does it require a parallel stack?

Maintained, not snapshotted. Are deployments, role changes, contract migrations, and terms amendments tracked on an ongoing basis with effective dating?

These are the criteria we would apply, and we would apply them to ourselves.

A closing word on where this comes from. Lukka builds and maintains the layer described here: the terms attached to tokenized instruments, the participants behind them, and the connective tissue back to the identifiers institutions already run on.

None of it was designed in the abstract. The data model has been through repeated iteration, enhancement, and expansion, driven by problems brought to us by some of the largest fund administrators and banks along with other participants across the ecosystem. Several of the positions argued for in this paper are held because an earlier version of them did not survive contact with a real client's holdings. The paper is an artifact of that work rather than a proposal drawn up in advance of it, and the examples throughout come from the same place.

The criteria above are the ones we would want a prospective client to hold us to, and coverage is the one to start with.

About Lukka

Founded in 2014, Lukka provides enterprise blockchain data and software solutions to financial institutions, exchanges, fund administrators, and government agencies. Its platform transforms raw on- and off-chain activity into audit-ready intelligence, powering accounting, compliance, risk, valuation, and reporting workflows across the digital asset ecosystem.

Lukka operates under AICPA SOC 1 Type II and SOC 2 Type II frameworks, delivering institutional-grade data and infrastructure for the next generation of finance.

Disclaimer

Written by Dan Huscher, Chief Operating Officer, Lukka, August 2026