S&P Global
Market Intelligence

The Hidden Architecture of Enterprise Data

Solving the Quiet Crisis of Disconnected Data

Across financial institutions and enterprises, data is multiplying faster, and across more fragmented sources than ever. But without consistent identifiers to connect records across systems, even basic questions such as who a company is, what it owns, and how it is changing become surprisingly difficult to answer. As the volume of messy data increases, so do the risks, gaps, and missed opportunities buried in the disconnect.

To learn more or to request a demo, visit spglobal.com/market-intelligence/en/solutions/cross-reference-services.

Data Everywhere. Insight Nowhere.

In a data-rich world, institutions are facing a growing paradox: more data, less clarity. Across financial services, asset management, insurance, and enterprise tech, data lives in silos. Identifiers are inconsistent, names are ambiguous, corporate hierarchies are incomplete, and the result is a quiet, persistent drag on productivity, insight, and compliance.

Institutions are grappling with an explosion of identifiers: CUSIPs, ISINs, LoanX IDs, Bloomberg IDs, LEIs, DUNS numbers, VAT numbers, internal codes, and other proprietary identifiers. Each dataset speaks its own language, creating fragmented environments and making even basic questions—like ***"Who is this company?"*** and ***"What else do we know about them?"***—surprisingly hard to answer.

Consider a global bank juggling millions of transactions, each with a separate identifier. Or a private equity firm tracking thousands of portfolio companies with varying naming conventions. Or ESG analysts trying to map data from dozens of providers, all using their own standards. Without a shared framework, it's often unclear which records refer to the same company.

Every institution has a version of this problem. The CRM full of duplicates. The ESG feed that doesn't match the financials. The risk model that breaks because a company changed its name, or no one noticed it had been acquired.

This isn't a new problem. It's the same one data teams have been patching for years—mismatched IDs, duplicate records, disconnected systems. But with more data coming in from more sources, faster than ever, those cracks have widened. What used to be a manageable cleanup job has become an enterprise-wide liability. And without a consistent way to connect it all, that data loses its power.

Understanding Identifiers: The DNA of Financial Data

In data management, an identifier is a unique code or label assigned to an entity, whether that's a company, financial instrument, individual, location, or transaction. Identifiers make it possible to catalogue, search, connect, and analyse information at scale. They are essential to building coherent, interoperable data systems.

In practice, identifiers are used to track:

- **Legal entities** (companies, funds, subsidiaries)
- **Financial instruments** (stocks, bonds, loans, and other instrument types)
- **People** (traders, executives, signatories)
- **Transactions** (invoices, payments, filings)
- **Locations** (country of headquarters, country of incorporation)

Every identifier has a unique format and a specific use case. Most are issued by official (NNA) or semi-official bodies (GLEIF), standards organizations, exchanges, or data vendors.

In global finance and business, the same company might appear in hundreds of different datasets. Names might vary—"Apple Inc." in one, "Apple Incorporated" in another, and "AAPL" elsewhere. Without a consistent, unique label, these entries are impossible to reliably match across systems.

Identifiers solve this by acting as unambiguous keys. They let computers (and humans) say with confidence, "This record refers to ***that*** company or instrument, or corporate structure."

For example:

- A trader uses an **ISIN** to ensure they're buying the correct bond.
- A compliance team uses a company's **LEI** to complete a regulatory report.
- A credit analyst uses a **S&P Ratings Org ID or Fitch ID** to determine a company's risk profile.
- An investment platform uses an **internal ID** to track performance metrics across tools.

How They Work Together (and why it's complicated)

In theory, identifiers are meant to unify data. A portfolio manager should be able to pull records from multiple systems and instantly know they refer to the same bond or company. But in practice, it's rarely that simple.

Many identifiers are context-specific:

- **LEIs** identify legal entities but not brands or business units.
- **DUNS numbers** may be missing for private or international firms.
- **ISINs** identify instruments, not the companies behind them.
- **Ticker** symbols often change and move from one exchange to another.
- And many systems still rely on internal or unofficial codes.

This means that while identifiers are essential for modern data operations, they're not always compatible out of the box. Organizations need a way to translate between them—mapping SEDOLs to ISINs to LEIs, Tax IDs, internal codes, and so on.

Many firms start with manual reconciliation or fuzzy matching tools. While fuzzy matching can handle simple cases, it struggles with messy, ambiguous, or incomplete data. And these traditional methods provide no insight into the reliability or accuracy of their matches, forcing analysts into costly manual reviews. In fact, analysts spend a huge portion of their time just cleaning and validating data. When faced with thousands or even millions of entity records—many of them poorly standardized—this manual effort becomes impractical.

A huge portion of analysts' time is spent simply validating and cleaning data.

As data complexity grows, so does the challenge. ESG reporting demands alignment across dozens of providers. Regulations like **MiFID II**, **Basel III**, and **Dodd-Frank** require precise, auditable entity resolution. And with **50,000+ M&A events annually**, company structures constantly change.

Global expansion adds another layer. Different jurisdictions use different ID schemes, and providers often rely on proprietary formats. Without robust cross-referencing, organizations face fragmentation, duplication, and error-prone manual processes.

CRM Data: A Test Case for Entity Governance

CRMs are one of the clearest examples of how fragmented data creates inefficiency. They are critical for powering prospecting, deal tracking, portfolio management, and vendor relationships, yet are notorious for duplicate accounts, inconsistent names, and incomplete records. When records are inconsistent, it becomes nearly impossible to trust the CRM as a single source of truth. By linking CRM records to unique identifiers, governance teams can establish cleaner, more reliable "golden copies" of accounts, contacts, and transactions. This not only improves day-to-day CRM use but also provides a stronger foundation for connected systems like finance, billing, procurement, and risk.

Take the Test! Enterprise Data Health Check

	Yes	No
1. Can your teams trace any entity-level report back to a unique ID?	☐	☐
2. Do you use more than one identifier system (e.g. ISIN, LEI, DUNS)?	☐	☐
3. Can your systems recognize the same company across financial and CRM data?	☐	☐
4. Is your identifier mapping updated automatically and regularly?	☐	☐
5. Are your corporate hierarchies version-controlled and auditable?	☐	☐
6. Does your analytics team trust the integrity of linked data sources?	☐	☐

If you answered "No" to any of these questions your organization may be at risk of data fragmentation, inefficiency, and compliance gaps. In the next section, discover how Cross Reference Services can help you overcome these issues by streamlining entity resolution and building a more resilient data foundation.

The Identifier Infrastructure 30 Years in the Making

S&P Global's Cross Reference Services is the result of over three decades of continuous development, an infrastructure purpose-built to solve one of the most entrenched challenges in enterprise data: connecting identifiers across fragmented systems, asset classes, and jurisdictions.

What began in the 1990s as a way to map a handful of security identifiers has grown into a global solution. Today, the system reconciles more than 60 types of identifiers across **95+ million financial instruments** and **32 million legal entities**. It draws on hundreds of trusted sources, including regulators, national numbering agencies, exchanges, and internal systems, and refreshes every 15 minutes.

At the center is the **Capital IQ Company ID**: a unique, persistent identifier that serves as an authoritative anchor for linking data across records, hierarchies, and products. Mapping any one of 60+ unique identifiers to the Capital IQ Company ID provides a seamless and stable reference point for consistent integration.

The infrastructure is built on three core components:

- **BECRS (Business Entity Cross Reference Service)**: Resolves entity-level identifiers like LEIs, DUNS, internal client IDs, ratings IDs, VAT numbers, and other Tax IDs. BECRS disambiguates naming inconsistencies and links subsidiaries to parents within verified hierarchies.
- **GICRS (Global Instrument Cross Reference Service)**: Maps instrument-level identifiers—such as ISINs, CUSIPs, SEDOLs, and Bloomberg FIGIs—to a unified instrument record. This allows institutions to connect pricing, ownership, ratings, and analytics with confidence.
- **ISCRS (Instrument Cross Reference Service)**: Provides Industry recognized Sector/Classification codes —such as NAICS, SIC, NACE, and more.

These modules work in concert to create a consistent identifier spine across datasets, platforms, and workflows. Each identifier relationship is timestamped, version-controlled, and fully auditable—supporting both operational needs and compliance standards.

< < < > > >

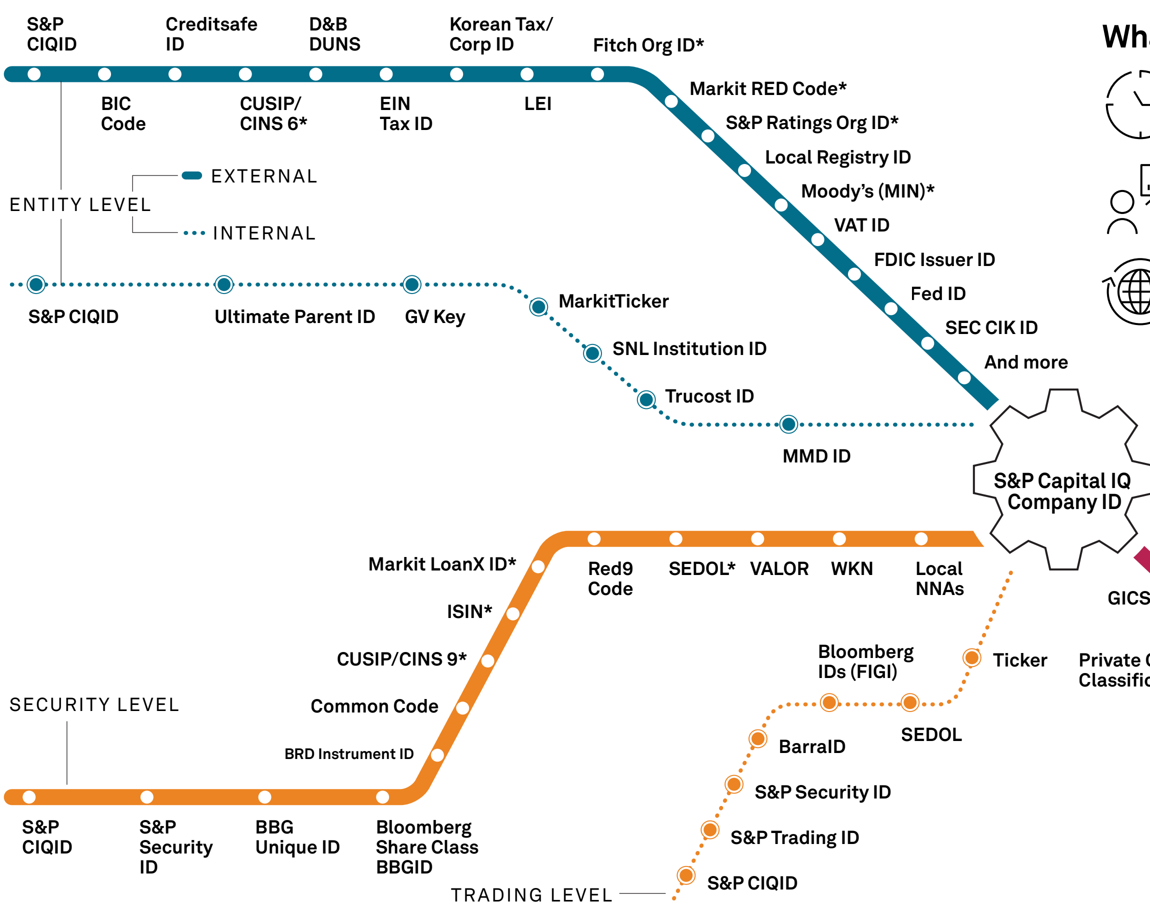
The seamless and stable reference point for connecting disparate identifiers

The S&P Capital IQ Company ID is a unique, persistent identifier that serves as the authoritative anchor for linking data across entities, instruments, and corporate structures, enabling consistent integration across systems and environments.

BECRS
Business Entity Cross Reference Services

GICRS
Global Instruments Cross Reference Services

ISCRS
Industry Sector Cross Reference Services



What's at stake?

- Faster Time to Insight
- Improved Risk and Compliance Accuracy
- Operational Efficiency
- Stronger Integration Across Systems
- Long-Term Data Resilience

*Requires a 3rd party license with vendor

< < < > > >

Beyond Fuzzy Matching — AI That Thinks Like an Analyst

While many firms still rely on fuzzy matching tools to reconcile names and records, these methods offer limited transparency and break down under complexity. Kensho Link is designed to solve the problem at its root, with machine learning models purpose-built to connect messy, fragmented data across companies, securities, people, and assets.

Using advanced natural language processing and large language models, Kensho Link automates over 70%* of the entity and record linkage process. It adapts to different data classes and linking policies, allowing teams to configure workflows with precision, whether cleaning CRM records, aligning ESG disclosures, or resolving legacy payment data.

Built with a modular architecture, Kensho Link offers:

- An intuitive interface for analysts to tailor matching logic
- Seamless integration into existing pipelines via API
- Transparent outputs, including match confidence and logic traceability
- Minimal need for human-in-the-loop review, thanks to secondary LLM-driven analysis

Unlike generic fuzzy matching tools, Kensho Link combines accuracy, adaptability, and automation, delivering cleaner data, faster decisions, and greater trust in downstream analytics.

*Based on internal assessments across client feeds and live production environments.

Fast, Accurate Company and Security Lookup

Not every workflow calls for bulk mapping or API integration. Sometimes the need is a quick, single, definitive answer. For quick, ad-hoc checks, The **Cross Reference Services Web Lookup Tool** brings the scale and precision of S&P Global's identifier infrastructure directly to the browser. It enables instant access to:

- 95 million securities searchable by identifier, company name, or security name
- 268 million security-level identifiers and 81 million issuer-level identifiers
- Verified issuer, security, and trading details, including ultimate parent company
- 122 million trading-level identifiers

[Get a free trial here]

What’s at stake?

Faster Time to Insight

Whether onboarding a new ESG data provider or integrating a portfolio of newly acquired companies, Cross Reference Services dramatically reduces time-to-value. What once took months due to manual matching, validation, and error-checking can now be done in days or even hours.

Improved Risk and Compliance Accuracy

With regulations like MiFID II, Basel III, and Dodd-Frank demanding precise entity-level traceability, having a trusted identifier spine is no longer optional. CRS and Kensho Link support audit-ready lineage, flag exceptions, and ensure consistency across internal and external reporting.

Operational Efficiency

Analysts spend a huge portion of their time cleaning and validating data. By automating entity matching and hierarchy resolution, firms can shift that effort from cleanup to value creation—supporting faster onboarding, sharper segmentation, and smarter decision-making.

Stronger Integration Across Systems

Linking all your data to the Capital IQ Unique ID provides a universal language for connecting data across silos—linking CRMs, compliance tools, data warehouses, ESG platforms, and trading systems to a common identifier framework. It reduces duplication, eliminates misalignment, and simplifies M&A data integration.

Long-Term Data Resilience

Data is constantly changing. Names evolve, identifiers expire, corporate actions reshape hierarchies. Cross Reference Services is built for this reality—continuously updated, version-controlled, and globally maintained by a team of experts. It ensures that today’s connections remain reliable tomorrow.

< < < > > >

Strategic Use Cases by Role

Compliance Officers

Ensure full entity traceability to meet MiFID II, Basel III, and KYC/AML requirements—backed by auditable identifier mapping and version control.

Data Engineers

Eliminate manual matching and fragile joins. Accelerate ETL workflows with persistent IDs and automated cross-referencing across systems.

Data Stewards

Maintain clean, standardized records across CRMs, warehouses, and third-party feeds. Prevent duplicates and outdated hierarchies at the source.

Analysts & Quants

Enhance model accuracy with consistently linked entities across time. Support better backtesting, valuation, and attribution analysis.

ESG Teams

Map fragmented ESG data to a unified company ID. Improve disclosure alignment, benchmarking, and impact tracking across providers.

Risk & Audit Teams

Trace exposure across subsidiaries and instruments with confidence. Support internal controls and stress testing with verified hierarchies.

M&A & Strategy Teams (Counterparty Risk)

Track acquisitions and restructures in real time. Link targets and comparables cleanly across internal and external datasets.

CRM & Sales Operations

Link CRM records to verified entities to eliminate duplicates, fix hierarchies, and improve segmentation and reporting.

Client Onboarding & KYC Teams

Speed up onboarding with validated identifiers and corporate hierarchies. Reduce friction in due diligence and entity verification.

Click here to request a meeting or visit our **website** for more information.

CONTACTS

Americas: +1 800 447 2273
Japan: +81 3 6262 1887
Asia-Pacific: +60 4 291 3600
Europe, Middle East, Africa: +44 (0) 134 432 8300

www.spglobal.com/marketintelligence
www.spglobal.com/en/enterprise/about/contact-us.html