Reference Data Acquisition Challenges: Getting it Right From the Start
TABLE OF CONTENTS
DATA ACQUISITION DRIVERS AND BENEFITS............................................................................................................. 3
DATA ACQUISITION CHALLENGES ......................................................................................................................... 5
DATA PROVIDERS .................................................................................................................................................. 5
DATA CONSUMERS ............................................................................................................................................... 7
CONCLUSIONS ...................................................................................................................................................... 10
ABOUT AITE GROUP ........................................................................................................................................... 11
CONTACT ............................................................................................................................................................... 11

LIST OF FIGURES
FIGURE 1: REFERENCE DATA MANAGEMENT KEY STEPS ....................................................................................... 3
FIGURE 2: REFERENCE DATA ACQUISITION KEY STEPS ...................................................................................... 4
FIGURE 3: FACTORS IMPACTING DATA VENDOR BUSINESS MODELS ................................................................. 7
FIGURE 4: MARKET DATA SPENDING INITIATIVES ............................................................................................... 8

LIST OF TABLES
TABLE A: REFERENCE DATA PROJECTS EXPAND IN SCOPE .................................................................................. 9
DATA ACQUISITION DRIVERS AND BENEFITS

Acquisition is the crucial first step in the processing of reference data among financial services institutions worldwide, including brokers, asset managers, and securities services firms (Figure 1).

Figure 1: Reference Data Management Key Steps

Without an organized, thorough, and consistent program to source, validate, and scrub data, the institution loses any benefits that a broader data management practice would provide, including the following:

- Improved risk management, including counterparty and liquidity analysis; this includes risk associated with complex listed and over-the-counter (OTC) trade processing and settlement operations

- Operations support, such as new security setup, which is especially important given that new instruments—including exotic derivatives, hybrids, and specialized/cross-category products—are almost constantly being introduced into the market to attract clients and boost revenue

- Data cost rationalization, a growing demand among firms seeking to cut costs

- Support for new regulations, especially upcoming Dodd-Frank Act provisions for centralized swap trading and clearing; buy-side and sell-side firms will experience implications including connectivity, straight-through processing, end-user support, record-keeping, and reporting


- Improved communication with customers and counterparties

Through conversations with reference data management professionals across the industry, Aite Group has identified key challenges associated with data acquisition. The following research, commissioned by Informatica, examines these challenges and the impact they have on financial institutions.

To put the issue in perspective, it’s useful to consider the scope of reference data acquisition. While the concept of acquisition suggests a fairly straightforward process, the reality is more complex (Figure 2). The process ranges across content analysis, requirements gathering, data mapping, business rule management, and maintenance.

Reference data can come from external vendors or be supplied via extraction from real-time data platforms that a firm already consumes. Many firms report that building the data loaders—the activity most commonly associated with data acquisition—is probably the easiest part of the process.

Figure 2: Reference Data Acquisition Key Steps

Source: Aite Group
DATA ACQUISITION CHALLENGES

Fundamentally, there are two components to the challenges most financial institutions face as they attempt to integrate reference data into their organizations: dealing with data vendors, and dealing with the various consumers of data inside of their organization.

DATA PROVIDERS

DIFFERENT SOURCES FOR DIFFERENT NEEDS

Many vendors provide data to financial services institutions. While the top two or three vendors have substantial market share, the industry often engages smaller vendors that have esoteric data sets. This, coupled with disparate needs from users, creates an avalanche of information. The multiple feeds support various business lines and span the front, middle, and back office.

Beyond the multitude of vendors supporting various trading and operational needs, many organizations implement data governance policies that stipulate the need for redundant feeds for audit and control purposes. This further magnifies the problem of data overload, challenging organizations to standardize their acquisition process to the highest degree possible.

DIFFERING LEVELS OF QUALITY AND TIMELINESS

Data vendors have their own data acquisition challenges to address, and they are daunting. Sourcing information from various exchanges, market participants, geographies, and asset classes can result in data of uneven quality and lacking in timeliness. Most vendors work to mitigate these issues, but their existence is to some extent a reality of the reference data marketplace. As a result, their clients must expend further effort to get the data into usable shape. In addition, feeds come in a variety of formats, including comma-separated value (CSV) and flat files, all of which must be managed.

One large, traditional asset management company employs 13 people in their data management group—half of which spend the majority of their time acting as “data stewards,” ensuring that data brought into the organization is scrubbed and validated, and any exceptions are dealt with. They sit among the business users and make sure, for instance, that any errors in particular data feeds are corrected in time to load into a trading system. This responsibility of data stewards—the need to vouch for data quality and support audit activities as needed—also creates additional overhead.

This process of error reduction is even more important in the context of composite, or “Golden Copy” records. A level of normalization must be achieved before attributes from different feeds can be brought together for a composite record. Due to a business users’ preferences with regard to a vendor’s data quality or timeliness, most firms conduct some level of composite record creation.
DATA FEED MAINTENANCE

Data vendors continually update their feeds to include new instruments introduced onto an exchange, for instance, or to support the additional attributes they collect.

The data acquisition team has to stand ready to collect notice of these changes, and then adjust their mappings when the time comes. This effort represents a recurring annual expense that will vary based upon the scope of reference data (i.e., cross-asset class or global data) being brought in. Doing so is an ongoing time commitment—these changes are frequent enough to necessitate nearly constant attention at some level, since notice of changes are released almost daily, with varying timelines for implementation. These types of operational challenges send firms toward a business-driven method of maintaining data feeds. A critical requirement includes enabling the business to easily maintain these feeds without IT involvement.

AUDIT

The point at which data is acquired and processed is a natural one at which to store the information in its raw format; it should be stored for potential audits from the data vendor to ensure the data is being used as laid out in contracts and for any regulatory or client audits that may come up. This is another level of activity among data management groups acquiring reference data.

In addition, if the incoming feed should cause an error, a report (to show the elements within the data that may have caused the error) and exceptions-handling process must be in place. As a result, firms are looking for a way to aggregate, normalize, and cleanse the data sources in a flow that can be monitored and audited.

DATA VENDOR EVOLUTION

The process of acquiring data is contingent upon commercial and operational engagement with data vendors. As in many sectors of the financial services industry, these vendors are seeing changes in their environment. This will impact how the acquisition of data is conducted. Two of these changes are described in Figure 3.
The results of these efforts for clients of data vendors will be mixed. One positive development seems to be an increase in value-added services to support the delivery and cleansing of reference data as vendors strive to rationalize their internal systems to ensure better quality and more timely delivery. While a step in the right direction, the effort and cost to acquire reference data will not be immediately impacted, since there will continue to be a multitude of vendors with which to engage.

**DATA CONSUMERS**

**PRODUCT AND GEOGRAPHICAL COMPLEXITY**

As firms expand their activity into new asset classes and geographies, they are experiencing a need to leverage a wider variety of data. This ever-changing process creates special problems for data management professionals, who must not only find the best sources of data for these new asset classes and geographies, but incorporate these feeds into their established sourcing, validation, and scrubbing routines. Finally, the data management team needs to facilitate new security setup on the appropriate trading, risk, accounting, and settlement systems, enabling new products to be traded. Introducing new products occurs across the enterprise, of course—not just for new asset classes or geographies. As a result, data management groups must handle a constant stream of new security requests.
Recent Aite Group research helps illustrate the degree to which new asset classes and geographic focus are a priority for financial services firms (Figure 4). We believe the prioritization of new technologies is linked to the need for data that covers additional asset classes and geographies.

**Figure 4: Market Data Spending Initiatives**

<table>
<thead>
<tr>
<th>Q. Where are you increasing spend in 2011? (n=34)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data on new asset classes</td>
</tr>
<tr>
<td>Data and applications for risk management</td>
</tr>
<tr>
<td>Staffing: Business analysis/management</td>
</tr>
<tr>
<td>New data from different geographical regions</td>
</tr>
<tr>
<td>Data feeds</td>
</tr>
<tr>
<td>New technologies</td>
</tr>
</tbody>
</table>

*Source: Aite Group*

**Organizational Changes**

As the financial services industry regains its footing after the credit crisis, mergers-and-acquisitions activity will likely increase. This has proven to be a major challenge for data management professionals, who must attempt to rationalize the overlap in data feeds they support between two organizations while maintaining an expected level of service to the various business groups. In addition, many large firms suggest that data management teams are often charged with soliciting business lines to subscribe to their centralized data repositories as a means to rationalize the amount of data being consumed (and paid for). This is an especially difficult process in the context of a recent merger.

**Increasing Scope of Data Management Projects**

Reference data management has traditionally focused on support for a variety of data types— identifiers, entity, ratings, end-of-day pricing, and corporate actions—inside financial services firms. Terms and conditions, contract terms, standard settlement instructions (SSIs), positions, and transactions data are also supported to varying degrees. Additionally, more and more market data (time series, valuation, historical, and analytics) is being included in this process (Table A).
Expanded scope suggests more challenges for acquiring data, of course. Another aspect of increasing scope that challenges organizations are service-level agreements (SLAs) put in place between data management groups and the business lines they serve. One major asset manager had to retool its data acquisition processes and technology to include sources representing 400,000 securities. This was necessary to meet an SLA it had in place with a global trading team that needed the data loaded before the market opened, a requirement it had struggled to fulfill.

SLA requirements have also increased the frequency of loading data into source systems. A process that once occurred on a daily basis, it is now much more likely to occur several times on an intraday basis. The result is an emphasis on accelerating the acquisition process to support these ever-increasing SLAs.

**SUPPORTING DISTRIBUTION**

Firms suggest that the effort involved in sourcing, scrubbing, validating, and normalizing data is critical to supporting the effective distribution of that data. One large firm manages multiple reference data feeds on a centralized basis, but is challenged to support operational systems, such as order management and accounting, while contributing data to a data warehouse used for reporting purposes. Disparity in how the data will be used demands a rigorous acquisition process to stage the data for customized distribution activities.

---

**Table A: Reference Data Projects Expand in Scope**

<table>
<thead>
<tr>
<th>Identifiers (CUSIP, SEDOL, ISIN)</th>
<th>End-of-day prices</th>
<th>Corporate actions</th>
<th>Asset class-specific (i.e., coupon, maturity, payment frequency, accrual)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entity data (name, address, place of registration)</td>
<td>Terms and conditions</td>
<td>ISDA documentation</td>
<td>Standard settlement instructions</td>
</tr>
<tr>
<td>Ratings</td>
<td>Margin details</td>
<td>Collateral details</td>
<td>Valuations</td>
</tr>
<tr>
<td>Historical data</td>
<td>Positions</td>
<td>Time series data</td>
<td>Analytics</td>
</tr>
</tbody>
</table>

*Source: Aite Group*
CONCLUSIONS

Firms report spending up to US$5 million annually on the technology and people needed to facilitate the reference data acquisition process. Many more millions are spent on reference data content. Where does this money go?

Firms struggle with a mixed bag of centralized versus federated reference data acquisition activities, and must maintain a continual balance between ensuring they are sourcing sufficient data to support their business and fighting cost-cutting initiatives, especially with regard to redundant spending on data.

Once they’ve navigated the challenges associated with managing data vendor relationships and internal business line support, the real work begins. As more firms formalize the process of data acquisition, the full scope of activities becomes more readily apparent.

Broker/dealers and asset managers report that a mix of well-planned data governance policies, savvy reference data professionals versed in the intricacies of particular vendor feeds and downstream application requirements, and the right technology infrastructure for the job are required to meet the increasingly complex and volume-intensive requirements of reference data acquisition.
ABOUT AITE GROUP

Aite Group is an independent research and advisory firm focused on business, technology, and regulatory issues and their impact on the financial services industry. With expertise in banking, payments, securities & investments, and insurance, Aite Group’s analysts deliver comprehensive, actionable advice to key market participants in financial services. Headquartered in Boston with a presence in Chicago, New York, San Francisco, London, and Milan, Aite Group works with its clients as a partner, advisor, and catalyst, challenging their basic assumptions and ensuring they remain at the forefront of industry trends.

CONTACT

For more information on research and consulting services, please contact:

Aite Group Sales
+1.617.338.6050
sales@aitegroup.com

For all press and conference inquiries, please contact:

Patrick Kilhaney
+1.718.522.2524
pr@aitegroup.com

For all other inquiries, please contact:

info@aitegroup.com