Research Article

Instrument Reference Data Management System for financial Institutions

Mr. Vijay Govind Kadam and Dr. Mininath R. Bendre

Department of Computer Engineering Pravara Rural Engineering College Loni, Ahamadnagar, India

Abstract

Financial institutions need financial reference data to plan future investment strategy so that they can meet all standards set by global regulatory bodies. Reference data management solution will provides reliable and trusted data that can be flown throughout which will not only save millions of dollars but also improve risk management and customer satisfaction for investment gain. In Instrument Reference Data Management (IRDM) system we have enables financial institutions to manage the entire instrument reference data environment from vendor data rationalization to instrument reference data architecture, design, and integration, and from indexing to automated data cleansing and distribution.

Keywords: Reference data, Financial Instruments, Financial asset class, Reference data Management

Introduction

Financial institutions deal with numerous financial instruments, ranging from stocks and funds to derivatives to meet ever increasing demand of the global securities marketplace. They need to tackle huge amount of data to trade and to keep track of these financial instruments. IRDM solution helps financial service institutions rationalize the process of reference data consumption. It is design to consolidate, clean, govern and distribute these key business data objects across the enterprise. IRDM includes pre define extensible data models and access methods with powerful applications to manage quality and lifecycle of reference data. IRDM is about creating and maintaining golden copy of master data from raw financial instrumental data provided by financial information providers and publish the golden data to consumers. End users will receive accurate and real time data of financial instrument which help them taking decision for their future investment.

Literature Survey

A. Golden Source Enterprise Data Management

U.S. based company named Goldensource [5] is one of the leading Enterprise Data Management and service provider, recently started working on to create Instrument reference data product which is still in beta phase and not fully out for use to users. Goldensource provides lot of functional information related to multiple financial asset classes like „Fixed income", „Equity", and „Derivatives" on its portal which help to understand financial entities and designed the data model for various asset classes.

B. Investment Banking EDM (Enterprise Data Management)

Most investment banks have their own enterprise reference data management system [6] which they use for their internal purpose. Investment banks also share the functional information of stock market trading structure with entity classes traded on it which help to understand stock market related information where asset class like Equity, Fixed Income, and derivatives get traded.

Why IRDM

Financial institutions, Exchanges, and market participants are undergoing a fundamental transformation, data management is become increasingly challenging. In this context it is extremely important to manage creation and maintenance of data to ensure its relevance and mitigate risk arising out of data inconsistency [1]. Data accuracy and reliability is mission critical and key enabler for all business operation including trade execution, risk management, and compliance reporting. Data management is the development and execution of architecture, policies, practices, and procedures to manage the information lifecycle need of enterprise in an effective manner.

Effective data management calls for seamless integration between all elements of overall data management lifecycle strategy, governance, operation, review, analysis, and actions.

In the most financial institutions, data is spread across multiple regions, departments, and systems. Many of these entities have reference data pertaining to parent company. However, they are enable to do so with ease when there is no central source of data. Instead, entities have their own nomenclature and data source piled in silos, with redundant system designed to extract and process data for individual requirements. [1] Apart from being in and inefficient design, this is extremely cost ineffective and prone to data inconsistency. IRDM system address all the above stated issues. It is the methodology of managing the creation and maintenance of data that can be shared across multiple regions, departments, and systems. It collects data from multiple sources, normalize it into a standard format, validate the data for accuracy, and consolidate it into single consistence data copy for distribution. IRDM cannot be ignored anymore in financial institutions because of:

A. Verity of Financial Instruments

Brokers and dealers have created many innovative financial instrument products to attract the customers and investors. Currently there are more than nine million financial instruments trading on security exchanges in overall world markets, each need system to maintain timely, accurate, and detailed information. These new financial product and their complex structures are become challenge to the personals or systems who maintain this data for financial institution.

B. Changes In Stock Market Execution Process

Trading mechanisms in stock market have been changed by shifting composition of stock market participants. For example, there has been a rapid increase in number of hedge funds and the emergence of mega „buy-side" firm, many of which used program trading and other algorithmic execution models. Decimalization and program trading have led to reduction in trade size with corresponding increase in volume. These factors have put a strain on data management platform as they required to deliver high volume with low latency to black-box trading system.

C. Laws of Regulations and compliances

Many countries have placed their own Regulatory and Compliance laws on traders or financial organizations for risk analysis and management, also world trade organizations has their laws to mitigate risk. This has forced financial institutions to place high priority on creating accurate and timely data to feed internal risk management system. Now institutions have more responsibility on their shoulder. Wrong information can cause financial exposure and hefty fine on institution from regulatory bodies.

D. Data Providers/Vendors Dependency

Demand of industry for wide range of financial instrument attributes, which gives birth to entire sub-industry of vendors who specialize in asset class wise financial data capture and distribution. These vendors are playing increasingly significant role in managing and providing raw data of financial instrument. However managing multiple source of data providers is costly and inconsistent.

Proposed Methodology

Instrument reference data management system manage the instrument data to deliver a single, well define, accurate, relevant, complete, and consistence view of data across multiple regions, departments, and systems. Instrument data which need to manage that is used solely for relating data in database to information beyond the boundaries of enterprise. In Financial institutions, it includes descriptive information about securities, corporation, and individuals.
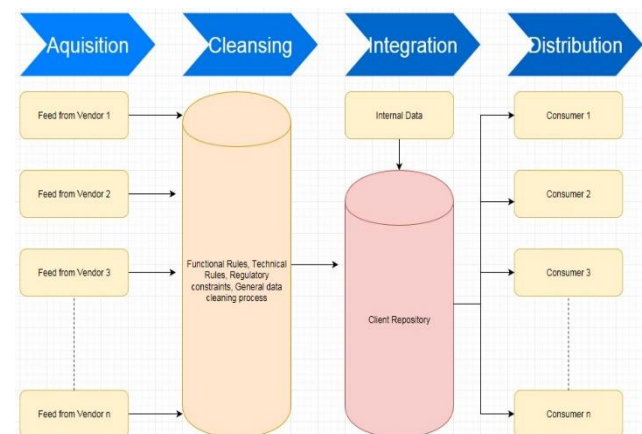

Fig. 1. IRDM Architecture

A. Proposed Architecture

Instrument RDM system is divided into four major parts namely Acquisition, Cleansing, integration, and distribution which covers life cycle of Instrument data management. Input to the system will be the raw instrument reference data provided by vendors, that input data will process in second layer by applying functional and technical rules. After filtering and cleaning data by rules, data will be stored in database through integration layer, and finally data will be distributed to consumers using distribution layer.
Acquisition: Instrument RDM system will take data as input from market data providers/Vendors who provides raw data from market transaction. Data from vendors can be in various format like JSON file, CSV file, xml format, and other formats. Instrument RDM system is capable to acquire input data in any format

provided by vendors. Input data can also be Table1. Sample Input data from Bloomberg provided on various destinations or locations like FTP server, web link, Web services, and on mail as per request. Instrument RMD is capable to acquire the data from such destinations.

Also input data can provided on specific time for example London stock exchange provides instrument trading data one hour after market closing time, Instrument RDM can also acquire such data automatically by creating schedule job for specific time to acquire data. Cleansing: After acquiring data from data providers/Vendors second major step of the system is to clean the data. There is separate module in system to clean the data to make it accurate and trust worthy.

Table1. Sample Input data from Bloomberg

SECURITIES|ERRORCODE|NUM FLDS|ID_BB_UNIQUE|ID_BB_COMPANY|ID_BB_SEC URITY|ID_BB_GLOBAL|ID_ISIN|ID_BB_SEC_NUM_D ES|FEED_SOURCE|MULTI_CPN_SCHEDULE_DT|MUL TI_CPN_SCHEDULE_PCT|MULTI_CPN_SCHEDULE_A MT|
BBG00H0FJQV6|0|10|COAO0489739|125919|1032113 71|
BBG00H0FJQV6|SG7DG2000002|WINGTAV4.08 PERP|BGN|20221228| |19325.3014|
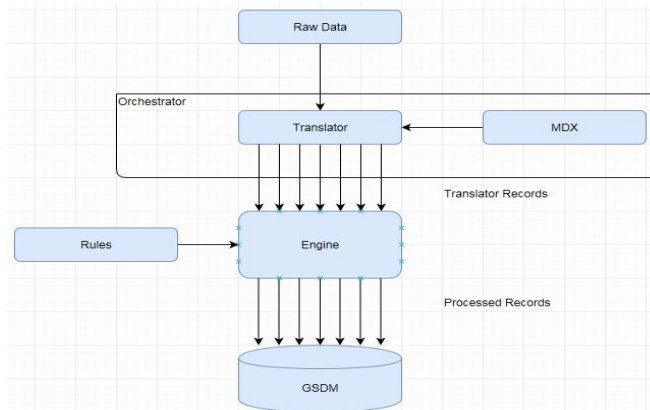

Fig. 2. Data cleaning module

Figure 2 shows the data cleaning process where three main part will work in coordination with each other namely Orchestrator, Engine and GSDM (General System Data management).
- Input: Raw data
- Translator/Interpreter: Orchestrator (Translator, MDX - Mapping Design Extensive)
- Processing: Engine (including Rules)
- Output: GSDM

Raw data will be received from vendors like Bloomberg, Reuters, Stock markets etc. that will be translated by orchestrator and converted into Street-ref file which will be xml file contents the segment of data mapped to specific table in database. Street-ref file

will be passed to Engine. For processing, Engine Will process street-ref-file by applying filters, rules that will make data accurate.

Integration: After cleaning, data will be stored in repository. Strongest part of this system is data model where cleaned data will be stored. Data model for this system is created in such way that all asset classes" data can be stored and maintain easily with BCNF normal form standard. If instruments data comes first time from vendor feed, which will be added in database, after that if any other vendor feed provide data on same instrument only changed data will be updated in database instead of updating all the data of instrument. Priority for vendor to enrich the data is also provided in system. For example instrument is added or enriched by Bloomberg vendor and user has given high priority to Bloomberg, then even if any other vendor provides data after Bloomberg that will not be updated in database as it doesn"t have higher priority than Blomberg vendor.

Table 2. Data Placed in Instrument Identifier Table

No	Instrument Identifier Table		
	Last change User Id	*Id context type*	*Issue ID*
1	BBCGCPFD	BBGLOBAL	BBG00H0FJQV6
2	BBCGCPFD	ISIN	SG7DG2000002
3	BBCGCPFD	SEDOL	BJQZLD1
4	CGS	CINS	V97973AJ6
5	REUITER	RUID	154599735

Distribution: Master copy of the Instrument reference data will be distribute the data consumers as per their need. Instrument reference data system will distribute data in various formats like XSL, JSON, XML, and CSV which are most common data format demanded by consumers. System will provide the data on destination like queue, FTP, web service response, and website link. System can also automatically mail the instrument data to consumers when data get updated from any input vendor feed.

Table 3. Data published to consumer in XML format

```
InstrumentHeader>
    <LocalIdentifier>76482648</LocalIdentifier>
    <InstrumentStatus>ACTIVE</InstrumentStatus>
    <ISIN>US749930AJ22</ISIN>
    <CUSIP>749930AJ2</CUSIP>
    <MarketSectorCode>Mtge</MarketSectorCode>
    <VersionNumber>38</VersionNumber>
    <assetClass>FixedIncome</assetClass>
    <assetType>MORTGAGE</assetType>
</InstrumentHeader>
```

B. Challenges

Improving data quality is an ongoing effort and financial institutions are facing the challenges of

improving their technology infrastructure to address this issue. Reference data management projects are major technology investments to improve the data quality. Data integration and concept of single source is a massive challenge as data is still being manage in silos. Some common challenges are listed below:

• Exponential increase in asset class, new securities and volume.
• Duplicate data vendor purchase, Expensive manual data cleaning, and poor data management leading to high aggregate cost.
• Management of multiple security masters, repositories, and different source of asset classes across different geographical markets.
• Prevalence to different identifiers like „Committee on Union Securities Identification Procedure"

cPGCON 2020 (Post Graduate Conference for Computer Engineering)
(CUSIP), „International Securities Identification
Numbers" (ISIN), „Stock Exchange Daily Official List" (SIDOL) and internal identifiers used by front and mid offices.

Result and Discussions

Reference Data is any kind of data that is used solely to categorize other data found in a database. In financial services, it includes descriptive information about securities, corporations, and individuals. Result of the system is to create and distribute master copy of Instrument reference data which is accurate and reliable. IRDM is capable to load Full file (All entries of Instruments) and delta files (Only difference of previous and todays reference data) in data model from vendor. IRDM is capable for continuously updated and monitoring of data as it is critical for successful data acquisition. Once data placed or enriched in data model it can be distributed to consumers as per their demand. IRDM provides audit trial and action tracking as it is extremely important for Error and performance management. Once we start feeding data to IRDM daily by receiving raw data from vendors, day by day IRDM data model will enrich with more accuracy and reliability of data.

A. Accuracy of Instrument Reference Data

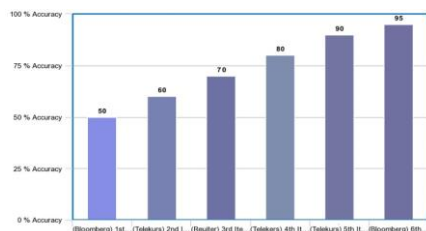Ultimate output of IRDM system is to provide accurate and reliable data to consumers.

Chart 1. Accuracy of Data over the Input Iterations

B. Vendor Reliability of Instrument Reference Data

If highest priority vendor has provided the data, then lowest vendor will not enrich the same instrument data. Vendor source hierarchy for asset class EQUITY is maintain like Bloomberg – Telekers – Reuter – Intex, then reliability of data from individual vendors are describe in below chart.
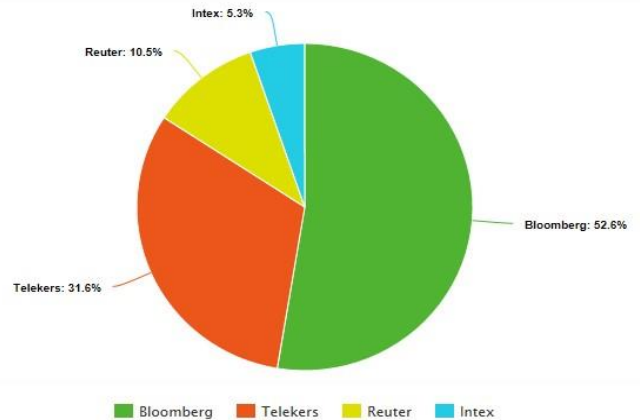

Chart 2. Vendor Data Reliability

Conclusion

Data across the platform is unbalanced, unstructured and not clean. If we want to use data for future investment in financial service we must have data balanced, structured and clean so based on this data companies, investment banks, Brokerage firms can use this this data for their future investment. Data must be acquire in raw format validate that data, clean it and then placed it in structure data model. IRDM is design and develop by keeping in mind the end users (Consumers of instrument reference data) those who are seeking to have instrument reference data to study and invest accordingly in future. IRDM simplify in four stages which helps system implement and integrate easily along with error and action management. In future all process of acquiring, validating, cleaning and distributing can be made automated. Data providers can provides data in any format or on any media our system should automatically process that data.

References

[1] "Master data management in financial services" – NIIT: White paper (2019) https://www.niit-tech.com/resource-library/white-papers/reference-datamanagement-financial-services
[2] Information Builder – White Paper: "Maximizing the benefit of customer centric master data management". (2018).
https://www.bigdatasummitcanada.com/wp-content/uploads/2019/02/WPMaximizingBenefitsOFCustomerFacingMasterD ataManagement2018.pdf [3]"Master data management" An Oracle white paper. (2011).
http://www.oracle.com/us/products/applications/master-datamanagement/018876.pdf.
[4] "Master data management – One step at a time". – White Paper (2006)
ftp://public.dhe.ibm.com/software/data/mdm/pdf/IDC_MDM.pdf.
[5] https://www.thegoldensource.com/resources/white-papers/.
[6] https://www.orchestranetworks.com/rd