MarkLogic®

ENABLING REGULATORY COMPLIANCE

This white paper addresses a key element in any GRC program: regulatory compliance. We discuss associated costs, current practices and risks, recommend a technologically enabled approach, and review the main business benefits for financial services, insurance and energy companies.

Contents

RISE IN BOARDS PREDICTING 'SIGNIFICANT INCREASE' IN COMPLIANCE SPEND FOR 2015-2016

YoY COMPLIANCE SPEND CONTINUES TO INCREASE IN THE US (USD)

MACQUARIE
+ 200% to USD 300 mln

HSBC
+300 to USD 750 mln

75% of compliance leaders expect management will require more/much more attention

Source: Thomson Reuters, 2015-2016 data.

EXECUTIVE SUMMARY

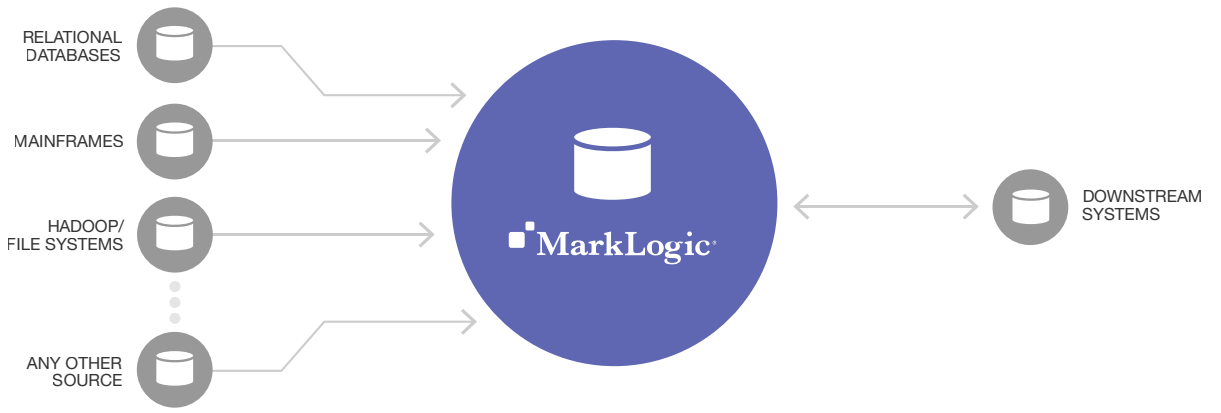
Senior executives in heavily regulated industries such as Energy, Financial Markets, and Insurance are constantly looking for optimization tools to enable an effective and dynamic Governance, Risk and Compliance (GRC) strategy. And, Compliance Management is a key component of that strategy.

In this whitepaper, we are looking beyond what compliance means and the costs associated with it. We identify key requirements and propose an effective technology-enabled approach to devising an effective compliance management system.

According to Thomson Reuters, compliance management is a holistic, connected methodology designed to protect and enhance business value by fostering a regulatory risk-aware culture. The graphic above shows some of the statistics quoted in their recent whitepaper "What's Compliance Worth?" to highlight the magnitude and importance of this activity.

All the compliance spending is typically only delivering a retrospective view – examining the operations after all the facts and figures are in. However, it is now an imperative to monitor processes in progress in real-time and proactively identify, even prevent, exceptions and violations.

Today, technology can enable the compliance function to run in-line with business operations by providing related functions across heavily regulated industries such as insurance, financial services and energy with the real-time feedback they need to make sure they are operating inbounds. Further, such technology is able to examine multiple lines of business simultaneously to get a "360-degree view" of every action taken and the possible relations and dependencies across these functions. Compliance can become cost-effective, agile and ultimately more proactive than ever before.



COMPLIANCE REQUIREMENTS

Fundamentally, the business requires a single platform of accurate data and timely analysis in an efficient manner. In a retrospective compliance model, the business wants to understand what is in the data they are providing the regulator – whereas what regulators and consultants are advocating is real-time monitoring and being able to flag anomalies and notify authorities before the process completes.

SINGLE PLATFORM

Enterprises are frequently fraught with disparate data sources, and analysis is frailly stitched together with a data warehouse which becomes a silo for another data warehouse in the future. Critical to the success of a compliance program is to have all of the data in a single place, ready to answer any question at a moment's notice.

ACCURATE DATA

Accurate data is a critical requirement in the compliance function. Data – both structured and unstructured – must be complete, reliable and consistent.

TIMELY ANALYSIS

Timely comprehension of data is critical in controlling costs associated with legal action. By providing the necessary data quickly, you can limit legal costs and potentially mitigate higher regulatory fines.

A RECOMMENDED APPROACH

MarkLogic(R) is a NoSQL multi-model database platform with a search engine and key enterprise features such as security, scalability, and ACID compliance built inside.

MarkLogic can be deployed as an *operational data hub* to reliably ingest and store any kind of data and make it instantly available to be searched, queried, and analyzed for compliance.

Implementing the MarkLogic capability will allow you to quickly answer and action upon such questions as:

- Did a new trade go against a current market trend or an analyst's recommendation?
- Does a change in the weather signal a need to re-examine hedging of a commodity trading strategy?
- Do the instant messages of a trader combined with their trading activity potentially indicate they are knowingly engaging in dishonest behavior?
- Based on what we know today, how may a trader's strategy have changed if they knew this yesterday?
- Would a claims adjuster modify a payments schedule based on new updates to the property & casualty database after assessment by a disaster management agency? Are these underwriters involved in a money laundering network?
- Is the provenance of these premiums clear enough to guaranty they are not part of a terrorism funding system?
- Are these transfer operations likely fraudulent, so that the company should not execute them?
- Should the company refuse to execute this withdrawal due to the receipt of a suspicious document?

These are just a few examples of the questions a Compliance Officer in an energy, financial or insurance firm may need to answer quickly.

MarkLogic gives you the ability to store all of the data you need in one place along with a suite of capabilities to efficiently answer these questions as new data comes into the system.

AUTOMATED INGESTION, INDEXING AND ALERTING

MarkLogic customers routinely ingest data into MarkLogic for compliance purposes as it comes into source systems. Customers who have created a "landing zone," "dumping ground," or "data lake" for their data can also use MarkLogic to aggregate data from these data lakes – making it easier to understand what kind of data you may have there, which is a key requirement for compliance. As data lands in the landing zone, processes can be put in place to automatically kick off an ingestion tool (such as MarkLogic's own mlcp) to ingest it into MarkLogic.

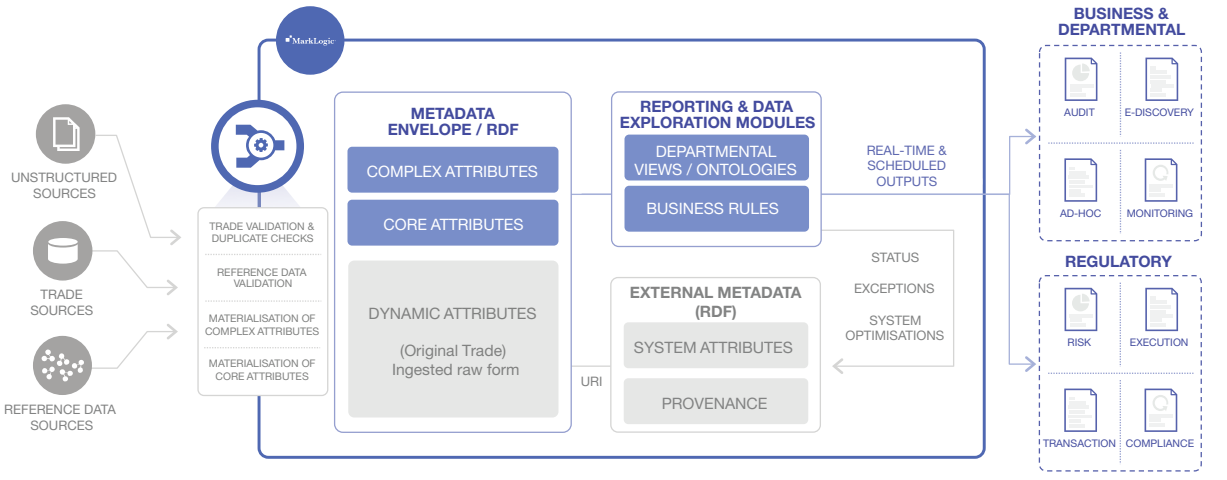
As the data goes into MarkLogic, additional processes can support data enrichment and molding into a form that is optimal for analysis. Further, alerts over email and other mediums may be issued if any "red flags" are identified within incoming data. For financial markets, this could mean ingesting trades, instant messages, emails, and various communication logs related to a transaction. For insurance, it could allow users to be warned that an awaited document has just arrived, that a customer requested help on social media, or that the amount of a claim reached a threshold that requires a supervisor approval.

Two main benefits of our approach are:

- Minimizing the time to apply compliance rules to data, without having to fully enrich or transform the data into its final searchable/queryable state. When something is in an upstream system or in a landing zone on a filesystem, it is generally "uncontrolled." MarkLogic allows you to categorise content and apply rules to it (e.g. compliance rules), by picking out the characteristics of the data that are relevant to the rules. That is done at the point of ingestion. Once something is committed to the database, it is controlled and compliant. You don't need a complete data model to pick out the bits that MarkLogic should use to apply compliance rules.

- Minimizing the time to allow search of data by compliance users. The MarkLogic Universal Index allows search across document content and structure as soon as the data is committed, without the need to do any upfront data modelling to define transforms into a normalized/canonical form. You can add indexes to specific things in order to make applications that use the data faster, but you can do that over time as you clarify/receive specific requirements from applications. You don't need to understand who the final set of users are going to be or what their requirements are, before you start loading in data. Your data model can evolve as your application does, saving your IT team time.

REGULATORY COMPLIANCE SOLUTION ARCHITECTURE FOR FINANCIAL MARKETS



SEMANTICS

MarkLogic is able to define related data after the data has been added to the database. Traditional data warehouses force you to define all the relationships up front. When a new relationship or view is discovered, the data must often be reorganized. MarkLogic is both a document store and a triple store, and uses semantics to reflect those relationships instead of traditional views and joins.

You can then ask the system to identify all the GE turbines in Texas and the system is able to infer that Turbine 123 is located in TX because it knows that Turbine 123 is in Abilene and Abilene is in TX. Or, the application can identify the taxation rules applicable to an investor A vs. investor B if they have the same asset in their respective portfolios and live in two different jurisdictions. The returns on the asset will be taxed correctly by inference as the recorded domicile of each investor would be sufficient to identify the proper tax policy. You also can use semantics to dynamically manage securities transparency by moving across defined characteristics and infer that a fund is composed of Asian shares, listed in Tokyo and Singapore, involving three different custodians and linked to industry and services, even though it is not described as such in the brochure of the issuer.

Semantics also allows your organization to build social graphs by linking and inferring relationships between people, locations, assets, and operations – and to visualize their evolution with a great depth of history. It helps, for instance, to better understand households' composition to target those where it is possible to sell more. It can alsol be used to detect, track and understand networks of fraud and criminal organizations using policies and claims to divert or launder funds, or to organize large-scale scams.

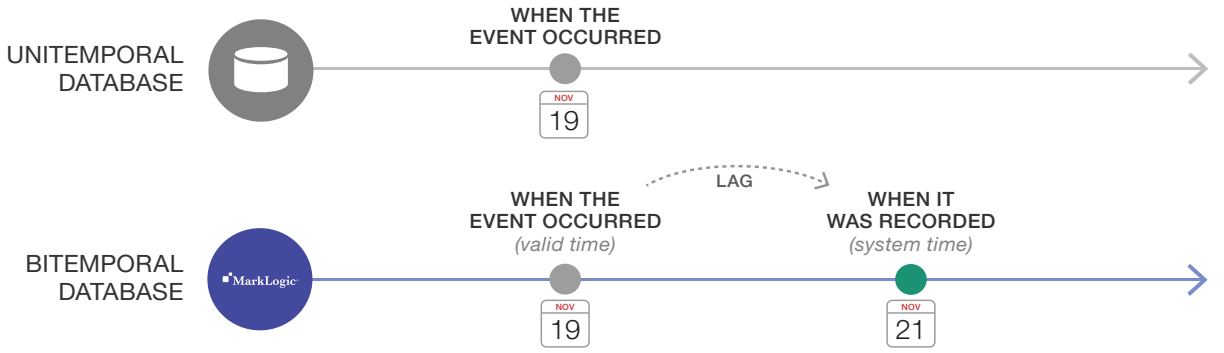
Although this sounds simple to you and me, it is a difficult thing for most systems to do. Semantics allows MarkLogic to define relationships much more naturally, using the way that humans think instead of forcing compliance professionals to think like software developers.

Semantics is great for allowing users to search data with "concept-based search" rather than looking for specific words or phrases. Thus a search can be done using purely business concepts such as underwriter, beneficiary, policy, claim or premium, regardless of the way this information is physically described and managed in different sources, thanks to collections which are logically regrouping multiple records and documents into common business objects.

BITEMPORALITY

We don't just look at things as they are happening. We also have the ability to look back in two different timelines using a feature called bitemporal time tracking. A bitemporal query can help answer "who knew what and when" by maintaining versions of the data as it changes over time.

Consider this example: a trader executes trade 1 and then executes a counter trade (trade 2) against the first. Another trade (trade 3) is executed based on the first two trades. If everything stayed intact then there would be no need for a bitemporal query, but sometimes trades are corrected or altered after the fact. If trade 2 was later changed, trade 3 and potentially trade 1 might look odd, perhaps even out of compliance.



Bitemporality allows you to look at the data as it exists now as well as how it existed at any point in the past. You would then know what it looked like originally and at every point between then and now. This can be done without special audit tables or consulting data backups. The system simply maintains the changed data alongside the current version. When a prior view is needed, the system brings up the version of the data that was in effect before the change.

This is particularly helpful when you want to retrace changes for instance on insurance policies when structures of the policy management system have changed, or the system by itself has been migrated, and retroactive operations have been done. In this case, with traditional relational systems, you would have to rebuild environments from backups and try to figure out how data was described and its current state, and then search across and query multiple environments with diverse data descriptions. With the MarkLogic bitemporal feature, you just have to run one query against a single environment.

SECURITY

Security is a feature baked into MarkLogic's core and was at the forethought of MarkLogic's earliest construction. MarkLogic offers the finest level of security granularity available in the NoSQL database market. Any piece of information or any function to access information can be secured through a variety of means with role based access controls to ensure that information is only seen by those who have permission to do so.

BUSINESS BENEFITS

Enabling a regulatory compliance function with the right technology can certainly result in tremendous benefits in the way the companies utilize and aggregate their data. We've identified a number of business benefits, below.

SELF-REPORTING

When mistakes do happen they can be caught early and self-reported, often reducing or avoiding fines. The ability to do this is especially key in the energy industry. FERC explains how self-reporting can allow a company to mitigate or completely avoid a fine:

The Commission encourages self-reports of possible violations, and in many cases, self-reported violations have resulted in closure of the matter without sanctions. In the cases where self-reports did result in enforcement action, the penalties reflected mitigation credit for the self-reporting that substantially lowered the amount.

Source:
http://www.ferc.gov/enforcement/self-reports.asp

Compliance professionals need to be able to construct a complete picture of an event, and salespeople and traders need to be able to see exactly what part of the behavior is in violation, while they still have time to correct it. Further, it's key that compliance shares only the data absolutely related to a violation without providing either too much or too little data to support a claim. Regulators will appreciate less data to review as well as being able to avoid sending subsequent requests for more data.

> The Commission encourages self-reports of possible violations, and in many cases, self-reported violations have resulted in closure of the matter without sanctions. In the cases where self-reports did result in enforcement action, the penalties reflected mitigation credit for the self-reporting that substantially lowered the amount.
>
> **Federal Energy Regulatory Commission**

Self-reporting has been difficult to do in the past because traditional IT systems are often rigid and don't provide the flexibility that a dynamic function like compliance requires. The old way of doing things was to gather information and run reports. But when the sources of information or the reporting requirements changed, it could take weeks or months before the compliance system could adjust.

The new compliance function requires an architecture that is flexible and scalable; one that can easily ingest both structured data like trades and records from commercial off-the-shelf software packages alongside unstructured data like emails, instant messages, voice and video, and documents.

QUICK AND ACCURATE REPORTING TO REGULATORS

Regulators by nature are curious. They look for trends and data points that fall outside of rational explanation. Traditional Business Intelligence reporting and data warehouses are highly structured and rigidly designed to ensure performance and consistency, but are built to answer only a defined set of queries. What is the cost to your organization to meet the second, third and fourth set of reporting requirements quickly – the ones you didn't know you were going to need to answer until the regulator asked for them? The modern compliance professional needs a system that allows for a plethora of disparate data to be ingested, preserved, and flexibly queried under one roof in order to answer unanticipated questions in real time.

AT A GLANCE SURVEILLANCE DASHBOARD

MarkLogic provides a powerful engine for storing and analyzing data, but also works with the latest web technology and reporting tools to provide powerful visualization of the data.

FACETED SEARCH AND QUERY

Faceted search is a form of guided navigation. Many popular websites have this type of navigational pane on that left that corresponds to categories that are dynamically generated using search terms. Dashboards with faceted search provide powerful insight into the business operations. Every piece of information that the system collects is indexed and available to our "ask anything" query engine.

CO-OCCURRENCE

Co-occurrence is a powerful query feature that can show you either values that co-occur frequently or infrequently, or allow you to find pieces of information where specific values co-occur. Imagine if you wanted to see which financial instruments co-occur with trading most frequently. For compliance purposes, you may want to watch out for a co-occurrence of a trader trading with a financial instrument they're not supposed to trade with. Or perhaps you're interested in seeing who (or which external company) co-occurs most frequently on the receiving end of a trader's communications. You can use this feature to analyze frequency of couples of values' occurrences, to detect unusual cases and determine whether they are normal or fraudulent. For instance, if a beneficiary who has never been involved in a policy appears as recipient

of a recurring withdrawal, this can be considered as suspicious and you can prevent validation by requiring a counter-validation, according to the “four eye control” principle. You can also use this feature to analyze loss notification and appraisers’ reports to determine co-occurrences of type of claims and type of damaged objects, age of insured and time of claims, car brands and loss valuation.

MarkLogic has a long and successful history of helping organizations simplify their compliance efforts across the world. We empower firms to integrate all of their data with minimal disruption to the business. MarkLogic’s multi-model database also provides powerful, “ask anything” search, semantic capabilities, and bitemporal stamping for querying transaction and valid times. With support for secure operations, accurate reporting and analysis over the full lifecycle of data, the agile MarkLogic platform makes it easier to adapt to changes – new regulations, governance and transparency standards – and innovate faster than ever before.

RELATED MATERIALS

Read our recent GRC blog to understand the building blocks of an efficient GRC strategy

Explore www.marklogic.com and find answers on how MarkLogic technology can help you expedite compliance.

MARKLOGIC CORPORATION
999 Skyway Road, Suite 200 San Carlos, CA 94070
+1 650 655 2300 | +1 877 992 8885 | www.marklogic.com | sales@marklogic.com

999 Skyway Road, Suite 200 San Carlos, CA 94070
+1 650 655 2300 | +1 877 992 8885
www.marklogic.com | sales@marklogic.com