

The status quo is changing

Data Recycling – Compliance Analytics in Times of Big Data

By Olly Salzmann



Before making a decision on whether to recycle an existing compliance application or invest in new big data technology, aspects such as cost, data quality and security have to be considered.

In the trough of disillusionment

The time for corporations to execute major investments on single-purpose tools only for compliance analytic services is coming to an end. In particular, the hype about eDiscovery tools and data fraud analytic services has entered the “trough of disillusionment” phase, according to Gartner’s Hype Cycle for Emerging Technology 2014 [<http://www.datanami.com/2014/08/20/hyped-tech->

big-data-iot/]. The focus remains on compliance, but the tools and analytics must also support the business in parallel. The key is finding the right “operational fit” for compliance in order to enable a better decisionmaking process for the business.

One solution could be to either recycle previously purchased compliance analytic services and tools and apply them to other business areas, e.g., enterprise con-

tent management, data archiving, financial exception reporting, supply chain process transparency, etc., or simply decommission/write off older investments and utilize data compliance analytic features already built in to the corporate infrastructure. The challenge will be finding the right balance between financial excellence and compliance with regulatory requirements. In the course of this article, two use cases will be presented, illustrating potential transformations of existing compliance analytics into advanced data analytics usable in both compliance and business contexts. The transformation can be achieved by either embedding existing compliance analytic tools into a new technological environment or migrating previously created data analytic functionalities to new big data tools, thus phasing out the old application.

How did we get there?

In general, the line between compliance-related analytics and business analytics, either from a service or tool perspective, has become blurred. Slowly but steadily, the traditional silos of data analytic tools and services ▶

used in either compliance or business will disappear. The common objective is getting more use out of existing data and simplifying the relationships and processes that come along with data. The prerequisite is the use of more advanced analytic methodologies along with a big data-supporting infrastructure.

In the mid-1950s, the first commercial analytics were applied to solve “shortest path problems” in the area of logistics. Since then, data analytics have steadily gained more attention, and thus evolved alongside the development of enterprise resource planning (ERP) systems, data warehouses and programming languages. Today, corporations of all sizes use data analytics.

At the beginning of the 21st century, various accounting and bribery scandals as well as the Zubulake case in 2003 and the amendments to the US Federal Rules of Civil Procedure in 2006 have led to a paradigm shift, increasing the importance of the role of compliance in companies. Approximately 20 years ago, the role of a Chief Compliance Officer (CCO) was rarely known, and if it existed at all, it was reporting to middle management. Today, the CCO is a high-power position, sometimes reporting directly to the CEO [<http://www.corporatecomplianceinsights.com/evolving-role-definition-history-of-cco-chief-compliance-officer/>]. The rise of more developed data analytic capabilities together with the focus on compliance topics within corporations have created a real hype at the beginning of the century. Therefore, compliance tools and data analytic services have evolved into a multibillion dollar market.

In the last couple of years the technology was changing tremendously, allowing for real-time processing, access to data that was previously not processable, distribution of low-cost storage systems (so called ‘data lakes’) and application of not just descriptive, but also predictive and even prescriptive data analytics. Topics such as Industry 4.0 and Internet of Things will have an impact on how regulatory bodies and compliance departments will operate. At the moment, more data is created every two days than the collective amount created between the beginning of time and 2003. From 2015 to 2020, data volume will increase from 3.2 to 40 zettabytes (40 * 109 terabyte or 40 trillion gigabytes), supported by the increase in the number of devices connected to the Internet from 13 to 50 billion [<https://www.linkedin.com/pulse/20140925030713-64875646-big-data-the-eye-opening-facts-everyone-should-know?trk=mp-author-card&trk=mp-author-card>]. Not only volume and the variety of data matter, value does as well. The likelihood of corporations to make decisions much faster than their competitors is five times higher when using big data analytics [<https://www.linkedin.com/pulse/big-data-amazing-numbers-2015-bernard-marr?trk=mp-reader-card>]. Predictive analytic capabilities could potentially be used to prevent negative business events from happening.

Business units and corporate functions within a company are becoming increasingly excited about the prospects that new technology and advanced analytics can offer. The era of big data could yield new business principles. Future competitive benefits may accrue to companies that can not only capture more and better data,

but also use it effectively across business and corporate units. Getting a handle on where to focus and how much to invest in or divest against opportunities and threats is becoming a crucial point on the agenda of senior executives.

Potentially two ways out

The buzz about big data and advanced analytics has been going on for a while, and there are many ways to fail, for early adopters in particular. Therefore, an objective assessment of the current as-is situation and subsequently, a solid transformation strategy is of the essence. The following usage cases present potential solutions to i) replace existing analytic features with more advanced features of a new technological infrastructure or ii) recycle already purchased compliance tools/analytics.

i. Vendor fraud versus master data management and procurement exception reporting

Traditional technologies, while still suitable and required for certain types of fraud prevention, are not necessarily designed to detect elaborate fraud patterns in the area of procurement. Uncovering procurement patterns with traditional relational database technologies requires a set of tables and columns and a series of complex queries. The tremendous effort and time required to write and execute these queries as well as the challenge to add and combine more data sources are huge disadvantages. In addition, the quality of how the query is ▶

written determines the relevance of the results. If exact questions are not known upfront, then the probability of discovering meaningful answers is close to zero.

Graph databases have emerged as an ideal tool for overcoming these hurdles. With the use of simple semantics, data can be analyzed, and new data sources can answer standard red flag tests, e.g., similar vendor names, duplicate bank accounts and addresses, as well as uncover complex data point correlations, i.e., combine text within emails or contracts and compare it with transactional data in payment systems). The database is able to use in-memory techniques so that processing in real-time is feasible. In combination with machine learning techniques, in which previously applied sampling methods and historical results are used to teach the system the relevance of results, a lot of false positives can be eliminated. Data visualization techniques will convert the generic table view of the results into interactive and correlated graphs, thus supporting the detection of suspicious patterns.

The new technological setup including the advanced analytic functionalities can be used to detect vendor fraud, but in parallel, provide additional reporting possibilities to the corporate performance controllers and supply chain management heads. The quality within the vendor master data can be monitored, and exceptions of procurement transactions can be identified. Obviously, the setup of a new technological infrastructure following the decommission of an existing application creates a lot of costs in the short term. Among others, the main cost drivers will be on data migration and se-

tup-inclusive license costs. However, in the mid and/or long term, a breakeven can be reached, especially as the new technological infrastructure serves multiple business areas and creates more transparency within corporate processes.

ii. eDiscovery for advanced archiving

In connection with keyword-based reviews of unstructured data in particular, e.g., emails and office documents, or structured data, e.g., accounting data, in the last couple of years, many companies have purchased an eDiscovery application, contracted an eDiscovery service provider or built their own application. The majority of corporations are using the – undoubtedly – very powerful data processing and text mining functionalities in the context of compliance investigations or other legal/regulatory proceedings. As of today, eDiscovery tools have been seen as a compliance enabler. However, the market entry of new and fast emerging technologies such as big data and advanced analytics changes the perspective of such tools from enabler to cost driver, at least temporarily until the new or updated technological setup has been found.

The technology trend goes to so called distributed file system solutions and in-memory databases, allowing high-performance access to data deployed on low-cost commodity hardware. Those systems are built to support applications with large sets of data whether structured or unstructured. The possibility of storing data on always-accessible storage volumes rather than archiving data to offline tapes enables an enterprise to query and

analyze historical company data in a fast and cost-efficient manner. This derives a competitive advantage and results in a better understanding of the own customers.

The data processing and text mining analytic capabilities within eDiscovery applications could make a good fit for processing and analyzing data in the context of enterprise content management and/or records management. eDiscovery data processing could filter, remove duplicates of and categorize data records before they get stored. Additional tracking and documentation features, built to be court-admissible, enable enterprises to comply with regulatory documentation requirements. Sophisticated text mining and tagging capabilities are a powerful toolkit to analyze and evaluate the content of the business records and support the creation and maintenance of a corporate memory.

The challenge in the short term will be connecting the frontend to a new backend and the big data infrastructure. The user-friendly and accepted frontend as well as the search algorithms and logic of the eDiscovery application could remain the same. However, the database and required storage location underneath the application must be changed in order to allow high-performance processing and access to a distributed low-cost file system/data lake. There are numerous technical solutions already available that can link a known application to a new technological big data infrastructure. But aspects such as security, disaster recovery and user access have to be considered among others before migrating this “recycled” setup to a productive IT environment of an enterprise. ▶

Conclusion

How much should an enterprise spend for a single-purpose compliance application and for how long? The current status quo is changing, and new initiatives such as Internet of Things, Industry 4.0 and big data are launched. This will have an impact on the type of data analytics that are executed for compliance, how they get executed, and mostly importantly, their cost. Ultimately, interconnectedness is growing, and traditional silos by function or department will be removed and turned into a more open and less organizational complex structure. Applications and their advanced analytic features can be used in various business areas and for different purposes. Therefore, single-purpose applications used only for compliance matters will have to be reevaluated and assessed to check whether the existing application and data shall be recycled or fully replaced by new and multi-purpose applications.

There are certainly new, promising opportunities available, but before making a decision on whether to recycle the existing compliance application landscape or invest in new big data technology, a few aspects such as cost, data quality, data privacy and security have to be considered. ◀



Olly Salzmann,
Partner, Corporate Finance,
Deloitte & Touche, Düsseldorf
osalzmann@deloitte.de
www.deloitte.com/de

www.businesslaw-magazine.com

Current Issue: June 11, 2015

Made in Germany

BusinessLawMagazine

Subscribe for free: www.businesslaw-magazine.com

The Business Law Magazine reports quarterly on all important questions related to German corporate, commercial, tax, labor, compliance and IP/IT law. Featuring articles written with real-world legal practice in mind, the online English-language magazine primarily targets company lawyers, managing directors, judges, prosecutors and attorneys in Germany's leading trade partners.



Published by



Strategic Partners

CLIFFORD
CHANCE

C/M/S/
Law. Tax

DLA PIPER

GIBSON DUNN

pwc

WHITE & CASE

Contact: FRANKFURT BUSINESS MEDIA, Karin Gangl, Telephone: +49 (0) 69-75 91-22 17, karin.gangl@frankfurt-bm.com