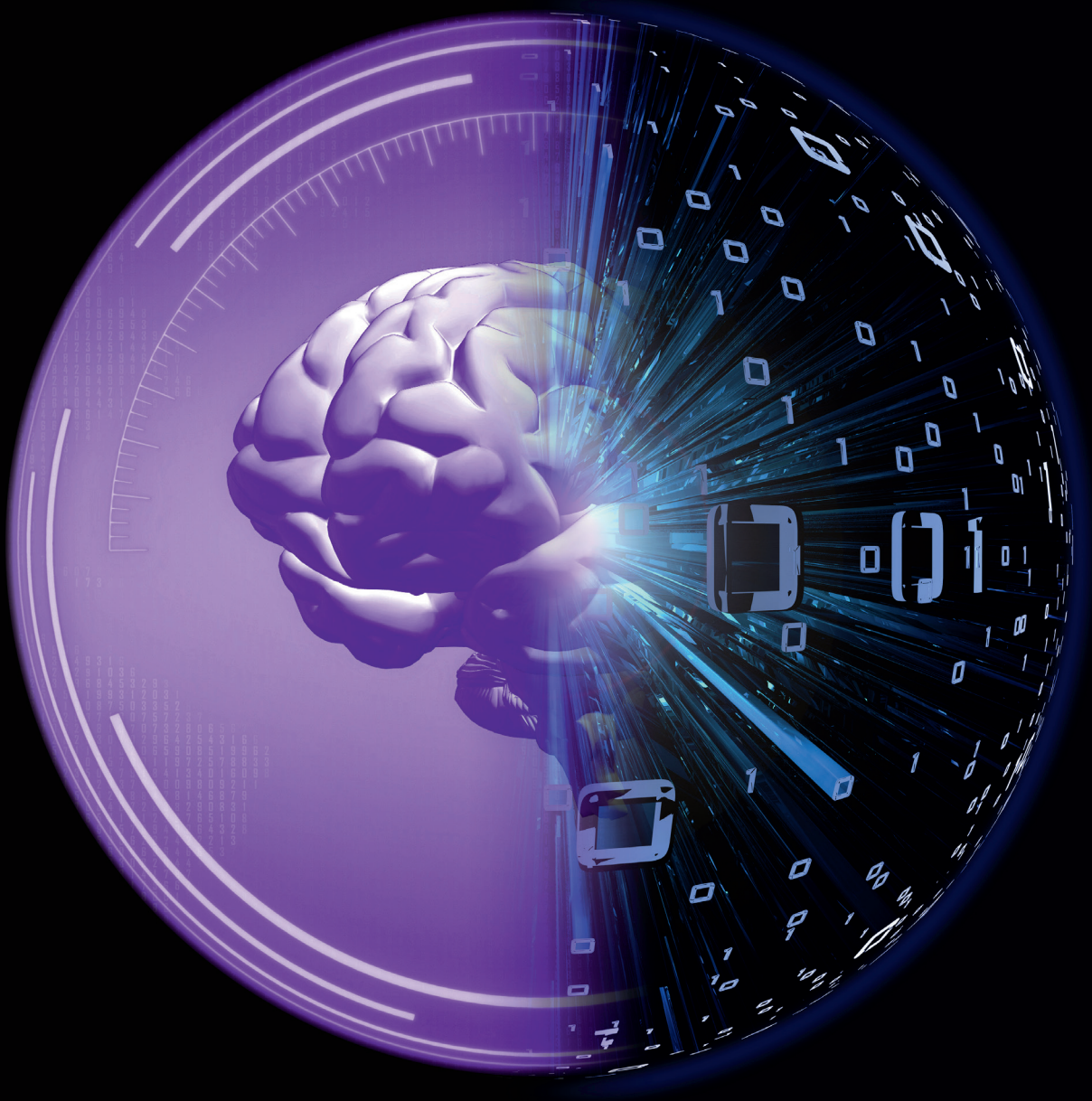
Deloitte.



The future of operational risk management

Evolving data architectures

January 2019

Moving beyond traditional operational risk data models to more integrated data structures for early risk identification, remediation and value creation.

In our paper *The future of operational risk in financial services*,[1] we highlighted how cost efficiency was becoming a higher priority in risk management and compliance. We also showed the consequent pressures on risk leaders to explore and embrace new technologies and techniques that can help improve the efficacy and effectiveness of their programs. We introduced concepts such as predictive risk intelligence and the use of advanced analytics for pattern recognition, as well as correlation and causal analysis to give operational risk managers a head start on identifying the buildup of potential risk and the need for remedial action.

Banks should seize the opportunities today's advanced tools and vast data pools make possible. Predictive risk analytics, machine learning, and artificial intelligence can help efficiently build and mine large and complex data sets that combine traditional Basel operational risk loss data with other data sources, including transaction data, non-transaction data (e.g., human resources, compliance, and other internal management information), and external data (e.g., sensing data, social media, customer complaints, and regulatory actions). These aggregated data sets provide billions of data combinations that can drive vastly improved analytical results and insights, and that can greatly increase the likelihood of uncovering patterns and correlations that previously weren't noticed until it was too late—if ever. This can help an organization prevent unpredictable outcomes and reduce operational losses and capital impacts.

Since our original publication in March 2018, we have seen only greater moves toward predictive risk intelligence. Globally, more banks are trying to make their operational risk management programs more forward looking. The purpose of this follow-up point of view is to highlight one of the implementation challenges to actualizing a more predictive operational risk management program. That challenge is the need for the evolution of the data architecture and models.

[1] Deloitte, The future of operational risk in financial services, available at – **https://www2.deloitte.com/us/en/pages/risk/articles/basel-final-rules-takeaways-highlights-us-banks.html.**

Note: While we use the term operational risk in this point of view, we recognize that some institutions have started to use the term Non-Financial Risk to include areas beyond the traditional Basel Committee definition (e.g., to include such risks as brand and reputation risk). Our definition of operational risk has always been such a broad definition, though we continue to use the terminology of operational risk.

Analytics strikes back – A case study in predicting patterns of sales fraud and misconduct

A global company was facing regulatory scrutiny related to its sales practices. It engaged Deloitte to help establish a proactive solution to detect likely cases of fraud and other sales practice risks. Deloitte designed an advanced analytics solution based upon rules, clustering, and predictive analytics to identify and forecast suspicious behaviors.

The solution linked data from various internal and external sources, including:

- **Customer account details**
- **Product snapshots**
- **Employee compensation**
- **Employment history**
- **Customer complaints**
- **Employee satisfaction surveys**
- **Employee reprimand records**

In addition to clustering techniques, the Deloitte team developed advanced analytics models to identify "behaviors of interest." For example, the team determined that one effective indicator of potential fraud was increased personal financial pressure on an individual, as represented by significant drops in variable income, combined with other faint signals, such as customer complaints or anomalous sales behavior.

The analytic solution ran on a periodic basis, and flagged potential misconduct for manual review by the client (in a manageable number of employees' activities, less than 0.5% of activity reviewed). The insights gained from this solution then helped the client make changes to their business processes. With Deloitte's assistance, the client developed rules and alerts to flag potentially suspicious sales activity across multiple behaviors and geographies. The Deloitte team then created customized dashboards and reports implemented as part of business as usual that delivered prioritized alerts to a client review team.

Most important to the client, they learned that their sales practices issues could have been identified up to two years earlier than they were, and they potentially could have avoided the regulatory fines and reputation damage they faced.

The phantom menace

Before we discuss what type of data model is relevant for operational risk, we should establish why data models are necessary in the first place. The main driver for careful design of most data models is to build a foundation that positions an organization to be able to derive better intelligence around a subject. Patterns and behaviors can help understand, manage, or predict the forces that drive them. Given the nature of operational risk, even predictable patterns and behaviors can still be challenging to identify consistently. Designing an adequate data model to manage this risk type is a challenge the industry has long known.

The initial constraint for the design of the historical operational risk data model was the singular objective of facilitating the estimation of conservative capital so the organization could absorb the impact of loss events. By design, this made it backward looking. The scope of data captured was narrowly focused on loss incidents. While this might have been appropriate at the time due to the risk of bank failures caused by operational risk events, the construct of early operational risk data models centered around inputs for mathematically modeling operational loss data to determine the adequate capital required to absorb such losses. There was little emphasis on more holistic and forward-looking risk management.

Significant challenges soon arose between what the models predicted and the reality of realized losses. This includes losses during and since the financial crisis. By their very nature, losses are realized with a lag, after a risk has materialized. In most cases, collecting only loss data doesn't provide assurance that all current risk exposures are identified. Historical data models did not comprehensively contain information related to all operational risk exposures, such as conduct risks, sales practices, and market manipulation—or the subsequent losses that could occur. Therein may lay the "phantom menace"—risks that are already materializing but with losses that haven't been recognized yet, and thus have not been captured in the data model or in the quantification of operational risk.

The nature of a loss can usually be attributable to the specific type of risk that has materialized. If the operational risk data model captures only losses that have arisen in the past, the model does not reflect the current risk exposure of the institution and potential future loss. In this age of rapid technological and business disruption, few organizations can confidently and credibly claim to capture that view.

The return—or rather continuity—of the GRC platforms

As we swiftly move into the new world of advanced analytic capabilities, the amount of data being added to the risk management process is vast. Operational risk managers, and other risk managers, have the opportunity to enrich the assessment of their company's risk exposure closer to real-time. Predictive analytics provides the opportunity to dramatically increase the quality of actionable insights. Risks may be more quickly mitigated as anomalies become more visible, against the context of the company's business objectives.

Governance, Risk and Compliance (GRC) platforms can continue to play an important role when insightful metrics, aligned with specific risk exposures and use cases, are collected and tagged with contextual reference data provided by a common taxonomy residing in a GRC platform. This information becomes an available node of intelligence for broader enterprise-level insights across various themes, as well as to inform risk assessments and mitigation strategies. When analytic results are placed into a GRC platforms' broader harmonized data model, they facilitate the connection to broader operational risk management themes, permitting the reporting across the range of the organization's risk taxonomy.

Leading GRC platforms can consume and organize metrics created outside the system. The functionality of a middleware aggregation and integration layer, combined with a data warehouse to support further analytics, potentially reduces the total cost of ownership by eliminating the task of building independent integration and aggregations/data stores. Another primary function of GRC platforms is to harmonize related taxonomy elements such as assessment units, business and functional processes, levels of risk definitions, controls, assets, obligations, policies and requirements, and other library elements. In a well-governed object data layer, risk and control ownership and accountability are well defined.

GRC platforms can also be a source of data for predictive analytic models. The "clean" data residing in the GRC platform (e.g., structured risk assessment scores and unstructured text comments) are pre-validated at the point of collection through agreed workflows with necessary reviews and approvals when data is captured. Current GRC platforms include integration (e.g., APIs) and data catalog capability where big data and "no-SQL" unstructured data may reside along with traditional structured data, making GRC platform data stores a good source of data for analytics models.

A new hope

So what is the solution? It might be to revisit the foundation of the operational risk data model—including the data we collect to identify patterns and behaviors.

One better way might be to learn from techniques derived from outside of risk management, such as customer marketing and sales. These disciplines have well-grounded techniques to help understand customer behavior to generate additional sales and further build customer loyalty. To derive those benefits, organizations had to monitor data from numerous sources so they could understand the full profile, preferences, and buying patterns of customer behavior. This ranged from monitoring and understanding customer traffic in retail institutions to developing merchandizing and designing websites and applications to increase sales and customer loyalty. In essence, this was a period of trial and error in understanding the customer interaction and engagement environment. Once built, it continues to evolve, adapt, and improve.

In operational risk management, we should emulate similar successes and begin to collect wide-ranging data through systems, applications, and processes—and through human interactions—then derive meaningful patterns and behaviors in line with the unique risk challenges of individual organizations and lines of business.

Only through the collection of this data at the broadest level can we identify patterns and behaviors and thus determine which data is truly risk-sensitive. We should look beyond losses if we hope to accurately determine the operational risk exposure of a firm.

Before we can accurately predict operational risk, we should first understand the relationship between the risk and data environment applicable to each business. Most institutions have taken shortcuts to identify metrics they deem to be risk-sensitive. But how do we know these are the metrics that are the most risk sensitive, if many are observed but not proven to predict? In reality, most of the metrics we call "key risk indicators" are developed following a significant risk or loss event, not derived or proven from an observed pattern or behavior.

In modern institutions—with the implementation of new technology, including robotics and process automation to replace manual processes —the availability of a wide range of data becomes far less challenging. This is where our foundational work around developing an operational risk data model should begin. Some may say that the current data environment is too vast and expansive to effectively monitor and evaluate. But with new ways to apply big data science techniques, institutions can now build these capabilities with relative ease and minimal investment. The real challenge will be in scoping what type and range of data will be relevant to derive the best model results. This is where leveraging business, as well as the experience of the operational risk manager, will continue to be key.

The figure on the right is an illustrative data architecture that highlights legacy Basel II components, those required for Standardized Management Approach (SMA), and a broader set of data sources required for predictive analysis. Broadly, this architecture includes:

(i) Data sources, which includes the systems interfaces, messaging, and data flows for bringing together currently disparate data;

(ii) Quantification calcul ators—the models that combine internal and external loss data to produce loss estimates (e.g., for current capital quantification and/or CCAR operational stress capital);

(iii) Core predictive analytics, to identify patterns, correlations, and causation that are otherwise hard to spot; and

(iv) Reporting capabilities—the mechanism for communicating current and potential operational risk exposures both to senior management and the business line units that manage operational risk on a daily basis, and integrate their feedback into traditional operational risk management processes.

While the structure is conceptually simple, there are several operational challenges to implementing this future state. Many organizations have faced some of these challenges while implementing significant regulatory programs that involve aggregation of data from multiple-sources systems.

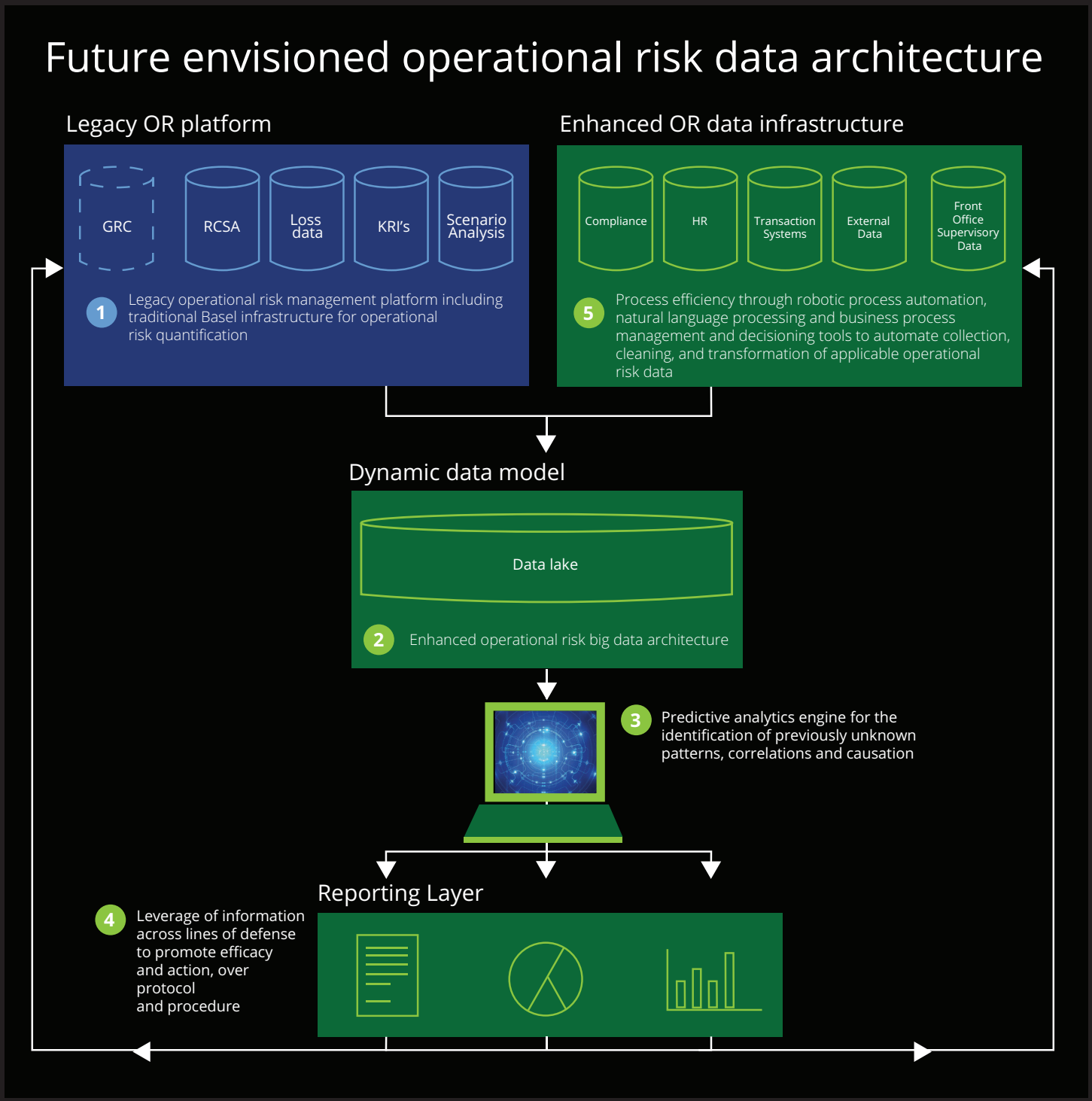
One key consideration in the case of operational risk analytics, however, is to refrain from creating oversized data pools in which the risk sensitivity of the data has not been established. This data risk sensitivity analysis is critical, because it will allow relationships to be generally pre-established in the dynamic operational risk model to improve risk detection and associated decision making. This is where the experience, judgment, and "smarts" of an operational risk manager can be the difference between boiling the ocean and collecting too much information, or progressing on this journey in a thoughtful, cost-effective manner that demonstrates quick wins and builds on that momentum. Stated differently, this new operational risk data model should be developed with defined rule sets that fuel deeper behavioral analysis, trend identification, and predictive analysis. Each organization will need to develop a unique set of characteristics and a bespoke implementation plan for a dynamic operational risk data model in line with its system and application architectures.

The data awakens

Operational risk data collected by organizations typically includes Risk and Control Self-Assessment (RCSA) results, internal operational risk incident descriptions and loss information, scenario analysis, issue management, and, occasionally, risk-oriented metrics.

There is, however, other equally valuable information that could inform operational risk managers but currently not collected—or if it is collected, it certainly isn't aggregated to provide a broader tapestry of the risk exposures the organization is exposed to. This information could include compliance metrics, front office supervisory data, HR information, and transactional data. The legacy data model tends to be less intuitive and predictive in effectively informing the organization as to measures, trends, and overarching risk profile.

Organizations now have the opportunity to expand the traditional operational risk data model. As organizations undergo digital transformations, the availability and range of data becomes easier to access and more readily available for consideration and potential inclusion in the newly defined operational risk data model. Moving toward a broader and more dynamic data model can open the door to more effective use of predictive risk analytics and allow data science techniques to assist organizations in understanding risk drivers, themes, and behaviors. The defining effect of these dynamic operational risk models can permit greater predictability and probability for organizations to determine their current level of risk.


Future envisioned operational risk data architecture
Legacy OR platform
GRC
RCSA
Loss data
KRI's
Scenario Analysis
1 Legacy operational risk management platform including traditional Basel infrastructure for operational risk quantification
Enhanced OR data infrastructure
Compliance
HR
Transaction Systems
External Data
Front Office Supervisory Data
5 Process efficiency through robotic process automation, natural language processing and business process management and decisioning tools to automate collection, cleaning, and transformation of applicable operational risk data
Dynamic data model
Data lake
2 Enhanced operational risk big data architecture
3 Predictive analytics engine for the identification of previously unknown patterns, correlations and causation
Reporting Layer
4 Leverage of information across lines of defense to promote efficacy and action, over protocol and procedure

Attack of the predictive analytics vendors

The predictive analytics market is served by multiple vendors—both established and nascent, emerging companies. While the solutions offered by various companies have some points of differentiation, most predictive analytics solutions offer some core features and capabilities, including support for data preparation and selection, insight generation, and visualization.

Most vendors support features for selecting data sources and formats and integration capability to seamlessly gain access to the data needed, evaluating multiple variables and selecting the ones to be included in the analysis, and ensuring the quality of data to be ingested—completing data sets, eliminating outliers, cleansing, and deciding what to do with missing values.

Many predictive analytics solutions also offer link analysis capabilities that can be used to visualize data to allow for better analysis. Link analysis has three primary purposes:

- **Find matches for known patterns of interests between linked objects**
- **Find anomalies by detecting violations of known patterns**
- **Find new patterns of interest (for example, in social networking and marketing and business intelligence)**

Vendors also offer predictive modeling capabilities that use data mining and probability to forecast outcomes. Each model is made up of many predictors, which are variables that are likely to influence future results. Once data has been collected for relevant predictors, a statistical model is formulated. The model may employ a simple linear equation, or it may be a complex neural network, mapped out by sophisticated software. As additional data becomes available, the statistical analysis model is validated or revised. Vendors are beginning to offer machine learning capabilities to help with the process of identifying the most appropriate (strongest) predictive model for a given data set.

Most vendors also offer embedded predictive analytics capabilities that can be used in the context of business processes. Embedded analytics can help organizations gain the visibility they need to understand current and historical results, as well as the causal factors influencing them. Embedded predictive analytics also allow organizations to predict system health and trigger alerts or to recommend corrective actions, helping ensure that systems are always performing optimally.

While most predictive analytics vendors offer the key features highlighted above, they differentiate themselves by offering additional capabilities in varying degrees, such as:

- **Ease of management: Unified platform, visual workflow design, ease of retraining models**
- **Advanced features: Automation of process such as data sourcing and preparation, text mining, advanced visualization capabilities, including interactive data views and reporting**
- **Integration: Capabilities to integrate with statistical programming languages, such as R and Python, support for multiple file formats, databases and data types, and open source innovation**
- **Training and customer support**

Summary

Many organizations have started on the journey to evolve their operational risk architectures. The data components and infrastructure that support operational risk are beginning to shift to include a broader definition of the relevant data elements, and predictive analytics and modeling. As operational risk management continues to mature, the future state is likely to look similar to what we have described in this paper.

Although many of the major losses in the last decade could arguably be attributed to operational failures, it is curious that operational risk management still struggles to carve out a permanent seat at the risk management table. To win that role, risk managers will need to demonstrate how operational risk management can help institutions meet their corporate and risk objectives by protecting their franchises and reputations. This will include the ability for operational risk managers to demonstrate that they are looking at risks the institution is currently facing, as well as looking forward to evolving and emerging risks, and designing the appropriate risk mitigation responses.

Contact us:

Monica O'Reilly
US Regulatory & Operations Risk Leader
Deloitte & Touche LLP
+1 415 783 5780
monoreilly@deloitte.com

Nitish Idnani
US Operational Risk Leader
Deloitte & Touche LLP
+1 212 436 2894
nidnani@deloitte.com

Steve Bhatti
Specialist Leader, Operational Risk
Deloitte & Touche LLP
+1 617 437 2451
stbhatti@deloitte.com

Kristen Gantt
Specialist Leader, Operational Risk
Deloitte & Touche LLP
+1 212 436 4161
kgantt@deloitte.com

Special thanks to the following contributors to this publication.

Neal Gregory, senior manager, Deloitte & Touche LLP; **Nelson Coutinho**, manager, Deloitte & Touche LLP; **Priyanka Pushkarna**, senior manager, Deloitte & Touche LLP; and **Smriti Jyoti**, solution advisor, Deloitte & Touche LLP.

Deloitte.

About Deloitte
As used in this document, "Deloitte" means Deloitte Tax LLP, a subsidiary of Deloitte LLP. Please see www.deloitte.com/us/about for a detailed description of our legal structure. Certain services may not be available to attest clients under the rules and regulations of public accounting.

This publication contains general information only and Deloitte is not, by means of this presentation, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional advisor.

Deloitte shall not be responsible for any loss sustained by any person who relies on this publication.