



Summary

“Big data, machine learning, business intelligence, predictive analytics and several other types of analytical activity serve this single purpose - to make more effective, efficient and timely decisions.”

The proclamation that data is our greatest latent asset is often made, and usually without much elaboration. Other than being a convenient cliché for technology suppliers, what does this assertion mean? To answer this question, we need to focus on a core business activity, executed thousands, if not millions of times in medium and large businesses every day. The activity is that of decision making, at operational, tactical and strategic levels within the organization. And since some degree of uncertainty accompanies all decision making, we can improve decisions by reducing risk through high-grade supporting information. For example, if we know that people in an age range and salary bracket are more likely to buy a product, then we could focus our sales resources on such a group. Better decisions, at all levels of the organization, represents the only legitimate use of the barrage of technologies that might be conveniently grouped under the heading of analytics. Big data, machine learning, business intelligence, predictive analytics and several other types of analytical activity serve this single purpose - to make more effective, efficient and timely decisions.

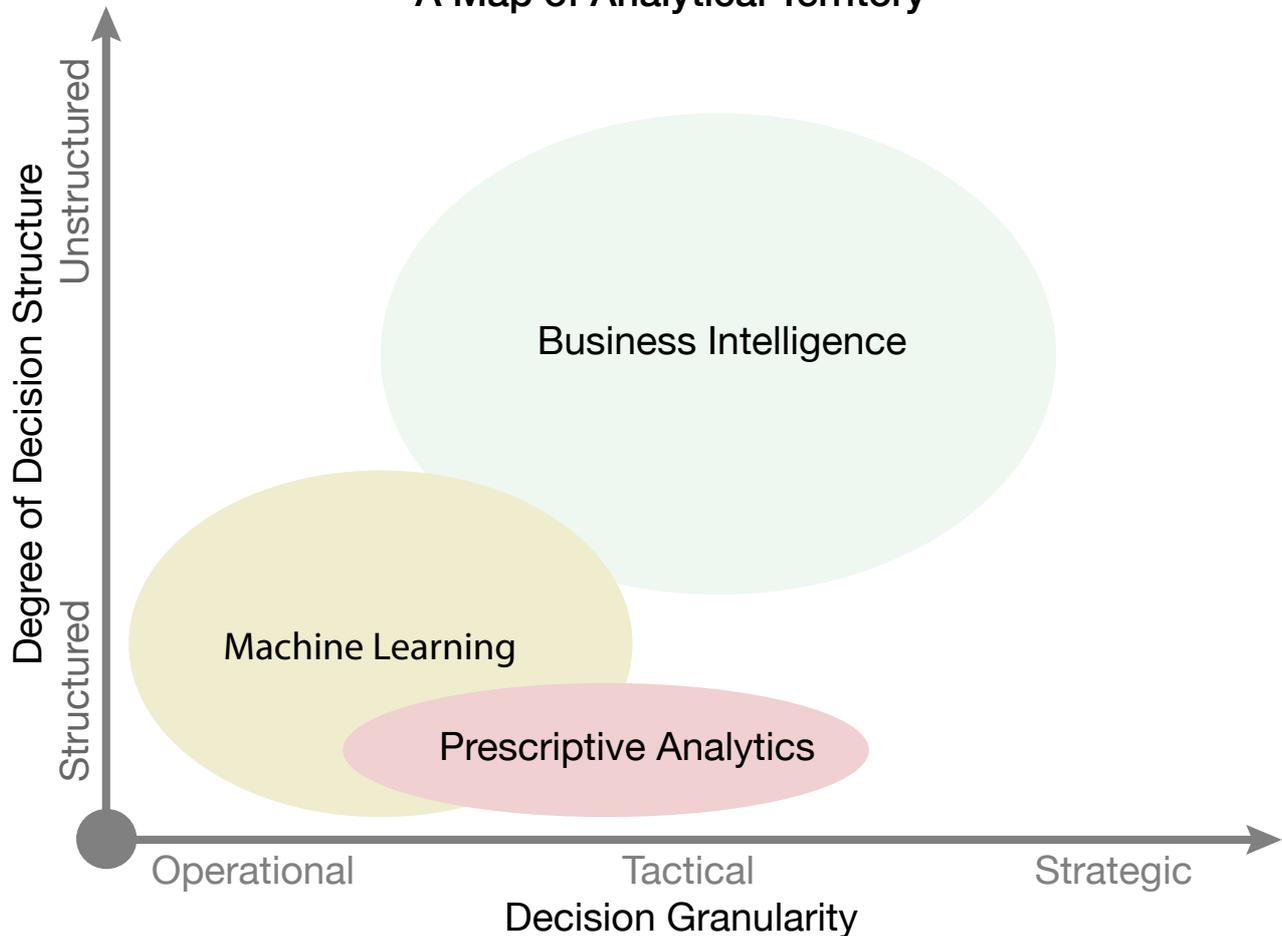
The operationalization of data is the act of garnering our data assets and processing them in such a way that they contribute to better decision making. Like all other investments, the cost savings and increased revenues that might derive from operationalizing data should be significantly higher than the costs associated with data, technology and skills. While this may sound obvious, the harsh reality is that many analytical activities are primarily experimental, without a clear relationship to operational decision making, and as such are a cost without any tangible benefits. Obviously, many organizations will need to engage in some level of data exploration and discovery, but that is not our concern here. We are interested in a direct relationship between operational efficiency and efficacy, and the exploitation of data assets.

Since most business managers understand very well the operational activity they manage, it is often quite straight forward to identify decision points where additional information might improve decision making. Examples of such decision points include loan approval, insurance policy pricing, identification of customers who might churn, possible fraud, deposit rate setting, recommendations and many others, depending on the business.

Decision making can be assisted or fully automated. Assisted decision making might employ various visuals displaying information in a readily digested format, or various recommendations derived from predictive models. In both cases it is imperative that the information is embedded into the production applications. Nothing is more harmful to productivity than context switching, and the embedding of information into commonly used applications completely eliminates the need to switch between applications on a frequent basis. Automated decision making is usually supervised in some way. Most applications for loans are now processed without human intervention, but even so people are still required to handle anomalies and marginal cases.



A Map of Analytical Territory



While the fact that the operationalization of data can, and does improve operational decision making is irrefutable, the execution of data operationalization involves a full life-cycle discipline. In essence this involves the integration of data sources, exploration of data, creation of analytical models (visual or algorithmic), deployment of those models in a production environment through embedding, monitoring of analytics performance from a business perspective, and retraining or modification of models as required. Clearly an integrated approach is necessary - an integrated technology platform, integration of activities between various departments, and integration down the command structure so the intentions of senior managers are realized in everyday operational activities.

The operationalization of data is no different from any technology investment and effort. Cost savings and revenue opportunities can be identified, and the relevant technologies deployed to realize them. Provided the whole effort is business driven, with a well-defined return on investment, businesses will be able to automate, partially or fully, yet another area of business operations. This process has proven profitable in process and transaction automation, and the same benefits will be realized in the automation of decision making.



Business Case

“Cost savings come from speeding-up decision-making processes, while more accurate decisions have the real potential to add to both top and bottom lines.”

The manual processing of operational decisions is a cost which can be reduced like any other cost. However, unlike many other systems and technologies investments, analytics can improve the effectiveness of decisions as well as reducing the cost of processing them. The cost reduction is simple enough to understand. In assisted decision-making information is provided at the point of work to speed up the decision-making process. Automated decisions naturally offer the most opportunity to reduce costs, but many businesses will want to apply analytics technologies incrementally, with the goal of eventually automating some of their decision-based processes.

Improving the effectiveness of decisions is perhaps the most significant part of the benefits which derive from analytics technologies. More accurate decisions may come from providing timely, accurate and relevant information at the point of work, in the form of visuals, or production applications may use embedded analytical models which recommend the best action to take. The lift obtained from such resources varies enormously, depending on the sophistication of the decision-making process before visuals and analytical models were deployed. Many businesses will be delighted with a lift of a few percent in reduced customer churn, while some companies will experience lifts far more significant than this - possibly in the tens of percent.

Automated decision-making employs decision models which have been trained using historical data. These models result from various algorithms analyzing historical data and finding useful patterns which can be used in current business activity. Such models have been shown repeatedly to be more accurate than manual processing, and this will inevitably be the way many decisions are executed in the future.

So, the business case for embedding analytics technologies into the production environment is quite compelling, and will only become more so. Cost savings come from speeding-up decision-making processes, while more accurate decisions have the real potential to add to both top and bottom lines.

The justification for investments in data operationalization and analytics is typically made on estimated cost savings. In reality, however, the driving force is often the desire to improve decision accuracy and target specific decision processes such as identifying customers who might churn, whether to extend credit to a particular customer, discount incentives and many others. Estimates of benefits which might be derived from increased decision accuracy are hard to make, but there is always some expected improvement that is assumed and this should be included in return on investment calculations. It is to be noted that this approach is a wholly different proposition from the notion that providing analytical tools to a broad swathe of employees will somehow deliver benefits. The advantages of such an approach are entirely



unknown before an investment is made, whereas data operationalization projects come with very well-defined objectives and the returns are much easier to estimate.

On the cost side of the equation, it is always best to use as many prefabricated hardware and software components as possible. This approach reduces time-to-market, is typically less expensive, and reduces the risks associated with frequent reinvention of the wheel. In practical terms this may mean adopting solution templates, using a proven hardware configuration and integrated software architecture. While many large organizations may need to set up bespoke hardware and software environments, these are typically used for prototypes and development and are not usually intended to support a production environment.

All-in-all the business benefits associated with data operationalization and analytics are relatively easy to identify. And it is also the case that we are rapidly moving beyond the 'Wild West Frontier' aspect of business analytics. The technology and technology configurations needed to provide operational analytics are now well understood, and so the risks are much lower than they were just a few years ago. There is also an experienced skill base and a population of suppliers who know what is needed to make operational analytics work.

Use Cases

“When analytics are close to the point of work and supporting operational decision making, there are several well-defined use cases which can increase decision effectiveness and efficiency.”

Most industries, from financial services to manufacturing, use analytics to improve their decision making. Traditional business intelligence is often the starting point, although this is more suited to tactical decision making, instead of the operational decisioning we are concerned with here. The main difference between tactical and operational decision making is the need for a “real-time” architecture, reliability, performance and elastic scalability of hardware resources to execute operational decisions. While traditional BI may address big-picture issues, it was never intended to serve as an operational decisioning platform. So the use cases we discuss here are specifically concerned with operations, and not the BI activities which revolve around diagnostics and description.

When analytics are close to the point of work and supporting operational decision making, there are several well-defined use cases which can increase decision effectiveness and efficiency:

- Passive visuals embedded into production applications. They are passive in the sense that they simply present information relevant to the task at hand, but include no added intelligence. A recruitment company might offer visuals to clients looking for certain skills, showing salaries for such skills in different cities, and the spread of salaries for similar skill sets. Such information is clearly useful, and can be considered as a starting point in the use of decision enhancing embedded analytics.



November 2017

- Active visuals include some level of added intelligence. This might be generated from the inclusion of data from predictive models, simulation algorithms, or prescriptive analytics where the best way of deploying resources is calculated. For example, a visual that supports simulation may suggest optimal human resources for a job, optimizing skills and labor costs. An insurance pricing visual might include data from a predictive model, showing the likelihood of a claim and a corresponding change of policy pricing.
- Embedded predictive analytics and machine learning may simply present additional fields on a form, suggesting a recommended action. A call center operative might be prompted to reduce a subscription rate to keep a customer who is likely to churn, or customer credit might be extended because a predictive model profiles the customer as being credit worthy. Recommendations are used widely in retail, and have applicability in most businesses.
- Automation of some decisions is an established practice in financial services, with loan approval and fraud detection being two prominent examples. Many loan applications are processed by an algorithm, and only spurious or marginal cases tend to get human attention. Fraud detection is also automated in credit and charge card businesses, and in some cases, it is even possible to detect well in advance whether a customer is likely to have difficulties with their repayments. Other industries also use automated decision making; good examples being the prediction of equipment failure in manufacturing, and automatic generation of offers to customers while in retail settings, based on their previous purchase history.

Financial Services

Credit and charge card operators make extensive use of analytics, and create data products that are used by partners, merchants, issuing banks and others. Fraud detection features very highly in the list of priorities, and card operators are always looking to minimize fraud while reducing false positives. This is difficult to achieve, and to fine tune the process a simulation dashboard might display the rates of true and false positives, and suggest modifications to parameters of the detection algorithm. It is also possible to detect customers who might have problems repaying their balance months ahead of any default taking place. As such, operators working for the card company can modify limits and engage customers well before problems occur. Recommendations can also be embedded into operational applications, suggesting other financial products a customer may be interested in. Such recommendations will typically come from a recommendation engine.

More generally, financial services and insurance companies make extensive use of all types of analytics to make better decisions. Insurance firms use analytics to help set policy prices, validate claims and target customers with other products. And as we've already mentioned, banks use embedded analytics to set deposit rates, loan rates and thresholds, and detect fraud. This all takes place against a background of strict regulation, and as such the analytical models must be transparent and well documented. Because of this the whole science and art of decision management tends to be more advanced in these firms than those in other industries.



Service Industries

Most service industries are essentially selling skills and solutions. To this end the use of analytics is often targeted at matching resources with need in the most cost and performance effective ways. Agencies who match skills requirements from large clients with their skills database will use dashboards and embedded intelligence to help engineer the best solution for a need. This might mean executing simulations that vary city, similar skills, combinations of skills and so on. In this way a client can be presented with various choices on location of skills and skill combinations so that needs are met at the lowest cost.

Other service businesses may be project based, needing to match skills and experience with project needs. Again, various dashboards can be used with embedded simulation so that the overall deployment of resources optimizes client satisfaction and profitability for the supplier. More generally the problem of resource deployment is one that is well addressed through prescriptive analytics, and particularly optimization. Embedding simulation and optimization into the operational environment can significantly increase the efficiency of many business.

Technology

*“The analytics layer is where intelligence is added to the data ...
Until data enter this layer, they are effectively dumb, with no added value.”*

The operationalization of data is achieved through the implementation of three technology layers: the data layer, the analytics layer and the presentation layer. Each of these layers might involve multiple technologies, but ultimately this three-layer architecture is what we need. The role of the data layer is to make data fit for analytical processing. This means it should provide mechanisms for data capture, storage, consolidation, transformation, and exploration and discovery. The technologies involved might include big data (Hadoop), streaming data, ETL, and data exploration tools.

The analytics layer is where intelligence is added to the data. Typically, this might involve predictive analytics, machine learning, data mining, prescriptive analytics, and possibly even artificial intelligence. Until data enter this layer, they are effectively dumb, with no added value. Most analytical methods involve analysis of historical data with the aim of discovering useful patterns that might have a use in current operations. For example, it might be found that people in a certain age range, and above a given salary threshold are much more likely to make loan repayments on time. The patterns that are found are often more complex than this, and finding patterns that are reliable is a skilled task. The intelligence contained in these patterns can be embedded into production applications in the form of recommended actions, or in some circumstances, they can completely automate a decision process. What is important is that the resulting analytical models are deployed at the point of work, improving decision efficiency and efficacy.



November 2017

The presentation layer is the one most familiar to people. The results of analysis are made available through various visual artefacts - charts, graphs, dials, or simply a number embedded into an application. Visual analytics has become very popular because of the immediacy of representation, and the fact that we process visual data very easily. However, it is often a mistake to believe that the visualization is the analysis, as is promoted by a large number of suppliers who operate in this space. The presentation layer is just that - a way of presenting the results of analysis, and not the analysis itself. Whenever possible the presentation of analysis should be at the point of work where decisions are being made.

Selecting a technology platform is complicated by the large number of options, and by the fact that the complexity associated with getting all the pieces to work together increases exponentially with their number. And so, it makes sense to employ an integrated platform where the interrelationships are known, and a single supplier is responsible for any compatibility issues. New developments in technology always start with technology and systems fragmentation, and eventually evolve into integrated platforms. Beyond this we find solutions being developed for standard business problems, and in analytics this might be customer churn, fraud, up-sell recommendations, and the like.

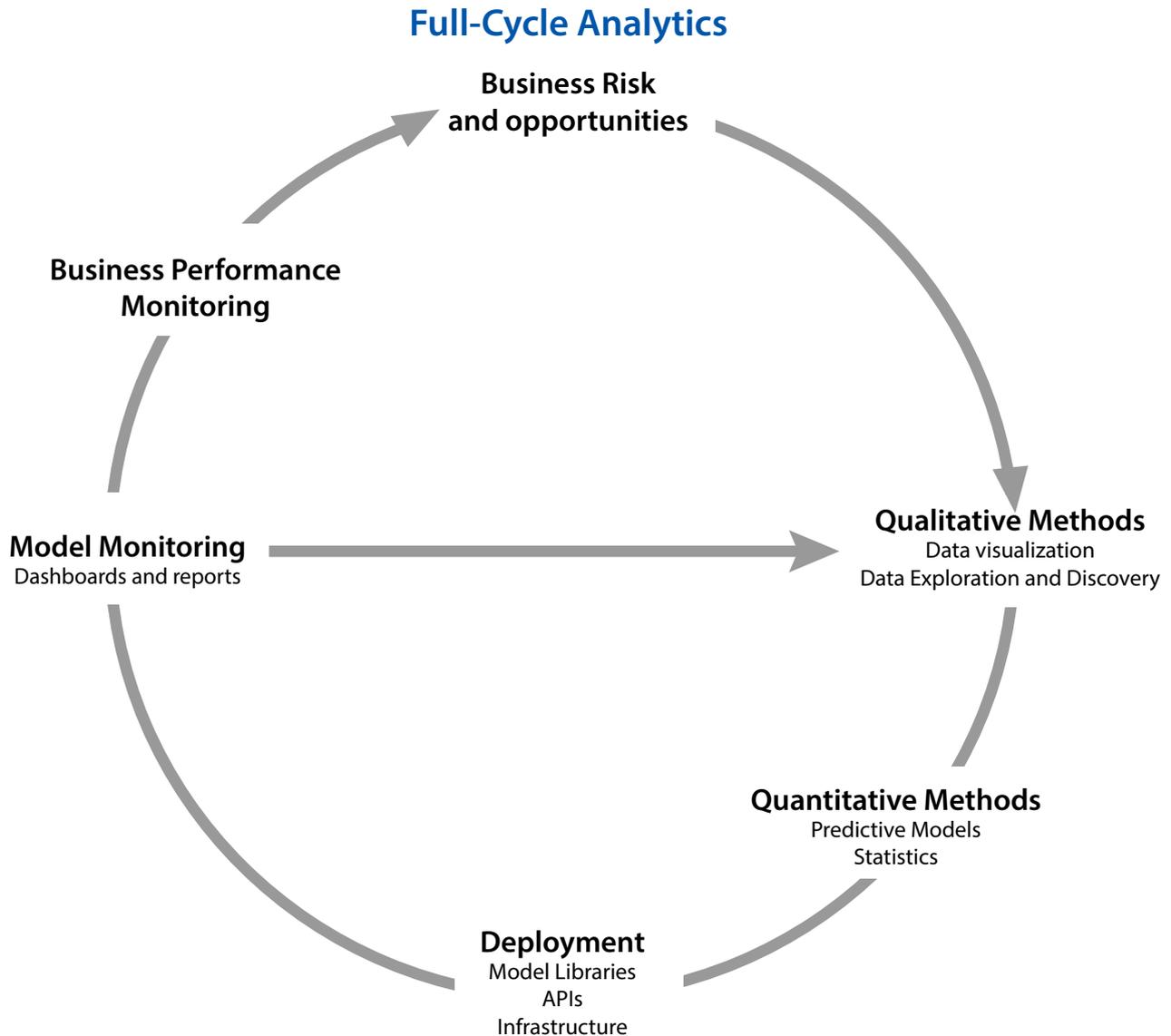
Governance

“The core task of governance in this context is to ensure that the lift obtained from an operational analytical model is maintained, and improved if possible.”

Since the operationalization of data primarily addresses the need for improved operational decision making, it clearly must be governed in some way. Accurate and efficient decisioning is the lifeblood of the organization, and so any systems which might enhance decision processes need to be strictly governed. Once again, we are not addressing ad-hoc and exploratory analytics here, but the analytics which are embedded into the production environment.

The development of decision enhancing analytical models begins with data. This is a whole topic in itself, but in outline data often need to be augmented by external sources, cleaned, checked for accuracy, profiled, merged and transformed. It is also necessary to explore data for the most meaningful combinations and transformations - a process generally known as feature engineering. This is an ongoing process aimed at increasing the value extracted from data.

The core task of governance in this context is to ensure that the lift obtained from an operational analytical model is maintained, and improved if possible. This can only be achieved if the tools are available to capture decision performance, and remedy any degradation.



In practical terms this means the data scientists need tools to interrogate the performance of analytical models, and business managers need business level dashboards to monitor how decision processes are performing. Such dashboards need to offer all the functionality we expect, and particularly the ability to drill down to details.

The governance of data operationalization, and the resulting analytical models it produces, involves several functions within the business. Data scientists and developers need an environment where models can be developed outside the production environment, and where model versioning is strictly controlled. The function responsible for deploying models also needs features such as the hot-swapping of models so that operational activity is not interrupted. Reporting between various functions also needs to be supported, so that if a business manager sees a fall-off in performance he can communicate this to data scientists and whoever else might be responsible for decision model quality.



Decision Management

“Full cycle analytics is at the heart of Decision Management, ensuring management can monitor and control decision processes, and that developers and data scientists can develop and modify decision models without interrupting operational activities.”

The effort to operationalize data and improve operational decision making has come under the umbrella term of Decision Management. As the name suggests this is a formal discipline in the same way business process management is. And since decision points in medium and large businesses may be numbered in the thousands, so it is necessary to set up various disciplines, adopt methodologies and deploy tools and technologies to design, develop and implement decision models. Of course, the models do not have to be implemented using computer systems, but in the main they will be since new technologies are quite capable of improving on manual decision-making processes.

Full cycle analytics is at the heart of Decision Management, ensuring management can monitor and control decision processes, and that developers and data scientists can develop and modify decision models without interrupting operational activities. In the same way a business process management system allows processes to be documented, designed, analyzed, deployed, modified, monitored and logged, so the operationalization of data through Decision Management disciplines must support the same activities. As such, governance is a core issue, becoming more important as the number of business decisions are enhanced and automated using the large array of technologies available to us. The alternative is uncontrolled deployment of decision processes that are undocumented and not monitored, with all that implies. And in industries that are heavily regulated, decision transparency is not optional, and so a Decision Management discipline is mandatory.

The Future

Assisted and automated decision making is already with us, and what is happening today is just a pale reflection of what will happen over the coming five to ten years. The emergence of artificial intelligence and machine learning has profound implications for the automation of many business decision processes, and with it the well reported disruption of working practices and white-collar employment. We've covered many of the issues in this paper – the adoption of Decision Management disciplines, governance, integrated technology architectures and embedding of analytics at the point of work. It also needs to be stressed that businesses will increasingly use solutions for their decision processes, in addition to bespoke analytical models.

AI based decision models based on machine learning, optimization, search and other techniques, already outperform human capabilities. Even the job of scanning x-ray, MRI and CT scans is now performed with greater accuracy by AI based agents, and some businesses



BUTLER

November 2017

From Data to Insights to Action

Make Better Business Decisions

have already started deploying AI ‘managers’ to implement scheduling and work allocation – Hitachi being one example.

Business intelligence as it currently exists will have disappeared within five years. The notion that dozens if not hundreds of people need to slice and dice data on a regular basis is fundamentally flawed in light of AI agents that can perform the same functions in just a fraction of the time. Several AI based BI products are already being used. These are not applicable to production analytics, but replace the labor intensive BI platforms typically used for tactical, and sometimes strategic analysis.

Finally, it needs to be pointed out that the technology standards are nearly all open source. Languages such as python and R account for nearly all analytical model creation. Big data technologies are almost all open source, as are the operating system components, and graphical libraries used to display charts and dashboards. The reasons for this are easy to understand. No organization can compete with the thousands of experts who contribute to open source technologies, and many of these experts are leaders in their field. Open Source is to a large extent future proof, and always sets technology trends. So, the future of decision automation technology is without a doubt Open Source, distributed through suppliers with the skills to provide integrated operational environments.

While the technologies are impressive it will as always be the skills of people that will determine whether assisted and automated decision making is successful. The stakes are very high in all of this, but so are the potential rewards.



Many thanks to **GoodData**, the sponsor of this paper.



BUTLER

Butler Analytics is a boutique technology analyst firm specializing in business intelligence, machine learning and artificial intelligence.