Integrated Enterprise Analytics

June 2015

Summary

Integrated enterprise analytics facilitates the rapid movement from insight to action. This is not usually a single step, but involves many steps, and initial insights gained from data exploration and discovery move on through more formal methods until something actionable is produced - usually a new insight as produced via some form of business rule or predictive model. An inability to actually operationalize that insight, while not killing such initiative, makes the exercise and technology investment useless.

Integration has several dimensions associated with it. Obviously it is desirable that the technology environment is integrated, in terms of both tools and infrastructure. But the sharper thorn is business integration. This typically does not happen unless forced, and the adoption of integrated Enterprise Resource Planning suites at the turn of the millennium (to mitigate millennium bug risks) imposed transaction integration on previously fragmented business operations. There are drivers for this in analytical applications, the most important being regulatory requirements. But the simple weight of analytical activities, with proliferating numbers of models, will force some level of integration - the alternative being analytical chaos.

The diagram shows a simplified schematic of the business analytics cycle. Analytics initiatives should always start with a recognition of business opportunity or threat - in other words, what is the business problem you are trying to solve. The next step is often qualitative in nature, using data exploration and discovery tools that are typically visual in nature. The discovery process gives insights into data, and an understanding that cannot be gained any other way. It allows business users to get a feel for important features in the data and their impact on the business issue at hand. The next phase is more quantitative and involves the use of statistics and/or machine learning methods. This has its own cycle of activities and is iterative in nature, as analysts and data scientists weed out erroneous models, and fine tune the features that have most impact. The next step is crucial, and in most real-world situations often suffers from delays, errors and dislocation. Given a well tested analytical model it is should be fairly straightforward to bring it into production. The obvious mechanisms are a business rules management facility, the creation of libraries of such models which
can be accessed as needed, a standard deployment language such as PMML (Predictive Model
Markup Language), and possibly an API interface. Most often however the model is handed over to a
programming team who might code the model in Java, or whichever language is used for such tasks.
This creates a discontinuity which in turn leads to maintenance issues as models are modified, and
the very real likelihood of errors. It is also an expensive way of doing things.

Once a model is in production it then becomes necessary to monitor its performance. An environment
which allows thresholds to be specified, and alerts to be generated is ideal for this purpose. Should a
model fail to perform well, it becomes necessary to revisit the quantitative analysis part of the cycle,
and maybe even the qualitative part. Finally, since business performance is directly related to the use
of analytical models, there is a need for a business performance monitoring, and ideally a means of
correlating performance with the deployment of models.

This is the ideal scenario, and one which represents a "Holy Grail" of business analytics. As such it is a
goal to be reached, and some organizations will be further down this path than others. And while we
have primarily focused on the technology aspects, this ideal involves organizational, cultural, and
conceptual changes in management thinking. The technology part of the equation, as always, is the
easiest part. Responding to commercial, regulatory, and organizational pressures is difficult because it
involves people. Nonetheless, integrated business analytics is where we are all headed, and necessity
will be the driver.

Integrated Analytics

An integrated analytics ecosystem would mean business intelligence, data exploration and discovery,
statistical analysis, machine learning and predictive analytics, text analytics, enterprise performance
management and other specialized analytic activities (graph analytics for example), would use
common data sources and have well defined interfaces. In reality these activities tend to be
conditioned by the genre of user, with qualitative analytical methods (data visualization for example)
and quantitative methods (statistics and machine learning) living in different worlds. A brief summary of
the different modes of analysis will be useful at this point:

- Business intelligence is nearly always used for diagnostic and descriptive analytics. It is these tools
  and methods which often inform of an opportunity or risk associated with current business activities.
  Scheduled reporting, dashboards and data visualizations are part of the repertoire of methods used,
  and always take a look in the rear-view mirror. Business intelligence forms the backbone of
  management intelligence in most organizations, and is closely linked with operational activities and
  data.

- Data discovery and exploration concerns itself with understanding data, and is more investigative
  than business intelligence. This is often a hybrid of quantitative and qualitative methods, and may
  use relatively simple constructs from statistics such as box plots, linear regression and other
  methods. This mode of analysis is relatively recent, and forms a link between qualitative and
  quantitative methods.
Predictive analytics and statistics are wholly quantitative in nature, although the resulting models may give insights into the nature of business operations. Predictive analytics often uses machine learning algorithms, which in turn need relevant features if they are to be useful. This is where the link between data discovery and exploration manifests, since it is these activities which inform business users, analysts, and data scientists on the nature of the data. As the name suggests, predictive analytics find patterns which can improve future business activities.

Enterprise performance management (EPM) platforms tend to inform management on the performance of the business. They typically use various dashboards with relevant KPIs with alerts and triggers to highlight exceptions. Business planning and formulation, financial management, strategy formulation and operational effectiveness monitoring all come under this umbrella. As the use of predictive models grows so it becomes necessary that these models and their performance is embraced by EPM. The alternative is that senior management is blind to the significant effect of predictive models, and have no way of incorporating their use into planning and strategy.

Finally there are several forms of analytics which are becoming more popular thanks to big data. Streaming analytics, where real-time events and patterns are detected, is exciting more interest as the Internet of Things (IoT) promises streaming data from devices and sensors. Although streaming data analysis, and particularly complex event processing (CEP) has been a focus in many firms involved in capital markets for some time. Text analytics, and particularly the detection of customer sentiment, is also a growing area of interest. This becomes more feasible with big data technologies, where document databases can handle large amounts of text data in a way that has not been possible previously. And as if this was not enough, there is also a rapid evolution of graph analytics taking place, where data with complex relationships can be analyzed at speed. Again big data technologies are accelerating this type of analytical activity, and it finds use in social network analysis and is used by many security agencies.

So it is not difficult to see that analytical activities are flourishing and proliferating. If we are to learn anything at all from the fifty years or so of transaction and process automation, we should already see that the analytics world is becoming fragmented, and with it comes duplication of effort, errors and inconsistencies, inefficiencies and isolation of business activities. Much can already be done to alleviate this situation, and particularly the use of an integrated Decision Management platform. There isn’t a single product available today that will integrate all the analytical modes outlined above, but some will integrate the majority. The integration of data visualization and exploration, predictive analytics, statistics, deployment into a production environment, and model management and governance is possible today. This heals many of the discontinuities which exist between business users, analysts and data scientists, programming teams and IT, and business management. What is more, suppliers who understand these issues (and many don’t) are the ones with the vision to eventually bring all analytical activities into an integrated whole. Integration is an inevitability, and more advanced users of analytics technologies are moving in this direction. The alternative is analytical chaos, harming the business and threatening reputations.
Customer Risk and Opportunity

It is no coincidence that the majority of business analytical activity centers around the customer. This is where most opportunity and risk is to be found, and so it makes sense to optimize these activities using analytical techniques and technologies.

Applications of customer focused analytical methods abound. On the opportunity side of the equation we want to offer customers the most relevant products at the most appropriate time. To this end analytical methods are used in marketing to focus campaigns on customers and prospects that are most likely to respond positively. It’s an all-round win situation, with campaign costs reduced (because fewer customers are targeted), reduction in the nuisance factor (communication is more relevant), and an increased lift, which in turn converts to greater profitability. Retailers particularly are using analytical methods to target customers more effectively. Location analytics has recently been added to the arsenal, where promotions are delivered to mobile devices when a customer is in range of a store. Offers can be made during the check-out process, and the positioning of goods in a store is very often driven by market basket analysis. Financial services also uses analytical techniques to optimize customer interactions, and because of the regulatory constraints, these same companies have to tie risk analysis in with their customer interface. And so loans are targeted at individuals whose risk profile is acceptable, and extensive use of fraud analytics means it is much harder for fraudulent intent to be successful. The rapidly developing Internet of Things IoT means devices and sensors give much more information about customer need and intention. The automobile industry is already placing sensors in vehicles which in turn can advise owners of an upcoming service, and this even extends to manufacturers of large construction machinery, where usage is monitored and advisories sent to users. Microsoft recently entered into an agreement with Miele, the white good manufacturer, so that streaming data from appliances can be processed, and various services offered to customers.

It should be clear that interaction with customers is only effective if it is targeted and timely. This implies two things - analytical models need to be developed for better targeting, and real-time infrastructure needs to be in place for timely customer interactions. However in many organizations these two needs are disjoint. The development and deployment of analytical models can be a lengthy process with various discontinuities slowing things down. One of the more serious is the discontinuity between model development and model deployment - a process that can be measured in weeks or months in many cases. Analysts and data scientists develop models, but it is often developers who code them and deploy them. Anything that can reduce this latency should be valued highly - although there are more than just technical and methodological issues at stake here.

Big data technologies mean we can analyze customers in many more dimensions. Graph analytics reveal relationships, social media analytics betray sentiment, streaming data can show real-time customer status, location analytics can deliver real-time opportunity - and so on.

Speed of model development and deployment needs to operate at a latency that is compatible with meaningful customer interactions. And there is a macro factor in all of this. Just as global
communications have increased economic and business volatility, so the ability to interact with customers in a real-time framework will also increase volatility. This race will be won by the speediest and most responsive organizations, and anything less than a fully integrated analytical environment will mean customers receive messages that are untimely and largely irrelevant.

**Actionable analytics**

As mentioned in the very beginning of this report – all meaningful analytical activities should lead to action. The various modes of analysis used by many businesses are currently mostly disjoint. Visual analytics, which is very much in vogue, is almost totally disjoint, and insights are often difficult to implement in a production environment. Ideally such insights should interface to the analytical models being used in an organization, after management consideration. In a sales situation for example, anomalies might be investigated using various data mining techniques to discover why they exist. Remedial action would then be possible by creating new predictive models and optimizing the deployment of resources.

In any viable system, action is only possible through integration. It is inconceivable that any entity, either animate or inanimate, could pursue meaningful action if its various parts were not integrated through meaningful communication. And so the insights of management need to be communicated to operations in a manner that is meaningful to both.

Business intelligence, data discovery and visualization, and data exploration are effectively the eyes and ears of the organization. Predictive models are the conditioned reflexes, which may or may not be appropriate in the current business environment, and for this reason need regular re-programming - or refreshes. This can only happen if lines of communication between the various departments in an organization are efficient, and a common language spoken. So in the example of sales anomalies, managers would talk with analysts to communicate the problem, and analysts could then set about establishing why the anomalies exist. Once the reasons become clear then some level of change in operational activities will be implied - unless the anomalies are accepted as inevitable.

Changes in business operations usually imply changes to business processes and business rules. Integration between these two aspects has recently moved forward with the development of Decision Model and Notation (DMN), an executable notation that injects decision logic into business processes. Change also implies a redeployment of resources in some way, and it is here that optimization techniques play a crucial role. In the sales anomaly example, it might mean pulling resources away from unprofitable customer segments and redeploying elsewhere. This type of resource allocation is enhanced considerably through the use of optimization, and in some real-world scenarios, such optimization might be pseudo-realtime as conditions change, and problems and opportunities arise.

The state of play today is that optimization and predictive modeling is integrated to some degree. Optimization can be used to determine which predictive models and business rules should be used in particular circumstances. Business processes can be integrated with business rules through DMN, which in turn can be used with business process modeling (BPM) methods and technologies.
Integration between the analytical modeling environment and production applications can also be achieved using standards such as PMML (although it does have limitations), libraries of predictive models and APIs. In reality however development and deployment of predictive models is also often disjoint - the models being reprogrammed in Java, or whichever language is used. The area where there is virtually no integration is between business intelligence (in all its forms), and model building. This situation is likely to get worse before it gets better, since there is a common misconception that visual analytics is a safe way to create business rules (making changes based on visual analytics implies some change to the rules driving operational activity). Until visual analytics is positioned firmly as a diagnostic tools and not a remedial one, then there will be little motivation to form a link between these two modes of analytical activity.

There is a long way to go in all of this, but it did take 50 years to automate business transactional activity and its management, and we are still in the early stages of understanding the avalanche of technologies which will inevitably automate many of our business decision processes.

Integrated Analytics Case Studies

While a great deal of focus currently rests of predictive analytics and machine learning, the resulting models are usually only part of a solution. Knowing when to utilize a particular model, and in which circumstances is also of crucial importance. The case studies shown below, courtesy of FICO the sponsor of this report, demonstrate the key role played by optimization in finding real-world integrated analytic solutions:

**A leading US retailer** wanted to leverage its extensive retail loyalty club program to improve the take-up and use of product offers. Its vision was to align its whole organization around a more member-centered approach. as part of its FICO Analytic Offer Manager solution, FICO generated thousands of propensity models at the product sub-category level for every offer available, which focus not just on what prospects are most likely to buy, but when they are most likely to make the purchase. The outputs from these are then used within FICO Xpress Optimization Suite to match offers to consumers in a manner that maximizes key performance indicators (KPIs) — such as maximize incremental sales that have a positive margin, whilst allowing for many member- and business-focused constraints.

The solution has provided positive results for the retailer across multiple campaign concepts and objectives, increasing response rates by factors of more than x5 on many segments. It has shown significant on-going benefit - Across different customer segments. - For both low and high tickets items. - And for new product trials.

**A large European Bank** aspired to implement a more customer-centric marketing campaign approach and needed an ultra-large-scale optimization solution to meet their business requirements. A decomposition technique, and use of the parallelization features of the FICO Xpress Optimization Suite, was used to solve the optimization model. This confirmed that the solution could run within the current infrastructure, on the whole customer base, within the required timeline. Additionally, the
Xpress-Insight module was used to quickly develop a series of scenario planning and comparison dashboards. The results of the proof-of-concept led to the full implementation of an integrated FICO Xpress Optimization Suite solution.

**A bank in the Asia-Pacific region** wanted to maximize profit on authorization & Credit Limit management decisions balancing between:

- Controlling accumulation of bad debt
- Attrition/customer service
- Revenue Generation

A decision optimization solution meant an additional $500K approved transactions in the first month, a 7% increase in income over five months with no increase in delinquency, and a projected additional $110M in transaction values over 12 months, with similar approval rates. A project for credit line increase realized $1.5MM in incremental profit in the first year, and the strategy was rolled out to 100% of the portfolio. Both projects created decision tree strategies for direct import into FICO TRIAD Customer Manager.

These examples illustrate well the advantages of combining various analytics techniques, something that needs an integrated analytic platform - in this case FICO.

**Integrated Analytics Strategy**

To integrate the analytics capability in an organization, there has to be something to integrate - obviously. The reality for many organizations, is that analytical activities are fragmented and partial. And since the ultimate goal of all analytical activities is to help automate decision making, either partially or completely, then this means that many decision making processes are still largely manual.

It is not uncommon to find business rules maintained manually in hard copy, and even then there will be modifications and exceptions that are undocumented and part of departmental folk law. In a busy business this is the line of least resistance, and so it is the one often chosen. Deploying business rules into a rules management system requires a level of intent and discipline that is hard to realize when day-to-day concerns are calling.

So, despite the illusion that businesses around the globe are automating their decision making by using both qualitative (data visualization) and quantitative (statistics, data mining) methods, the reality is somewhat different. And this is a good starting place!

To integrate analytical activities there are several bridges that need to be built:

1. Integrate qualitative methods with quantitative methods. Data visualization is a catalyst - the way we determine the state of play. Once anomalies have been identified, then we need the help of
qualitative methods to establish significance and possibly causes. Without this link data visualization simply reveals interesting events and phenomena, but no means of understanding them.

2. Quantitative methods reveal patterns that give insight into the behavior of customers, suppliers, machines - or whatever is of interest. Once we have such patterns we can exploit them by building predictive models and/or implementing rules within a rules management system.

3. Deciding when particular rules and models should be deployed for best overall advantage is an optimization problem. And so we use prescriptive analytics methods to optimize decision making.

4. The next bridge is crucial, and in many businesses barely exists. Deploying models into a production environment is best achieved using a rules management system when appropriate, a model library, a deployment language such as PMML, or possibly an API. In reality however, the logic embedded in models is re-coded in Java, C++, or whatever the language of choice may happen to be. This is fertile territory for errors, delays and unnecessary costs.

5. Once models are deployed they need to be monitored, and so the analytical toolkit should support some form of model monitoring and reporting. Some suppliers provide this level of capability, but many don’t.

6. Finally, it is clearly desirable that the effect of business rules and predictive models should be visible to management, via some form of enterprise performance management capability. This is almost entirely lacking in most businesses.

So in reality the integration problem can be clearly stated, and the actions needed to achieve integration fairly simple to understand. But simplicity does not imply ease, and realizing such an integrated capability will mean organizational, cultural and conceptual changes. This is always difficult and generally comes about through the force of necessity. It also implies that integrated enterprise analytics needs very high level sponsorship, unless analytical activity is to become fragmented and departmentalized.

There are big gains to be made from integrated analytics. Money saving efficiencies can be easily realized. The effectiveness of models and business rules increased. Room for errors decreased. Regulatory requirements are more easily met. And perhaps most important, management become engaged because they see the commercial impact of analytical activities. Integrated analytics is an inevitability - and some will get there before others.

About Butler Analytics

Butler Analytics is a boutique IT analyst firm specializing in business analytics technologies and methods. It was founded by Martin Butler, best known as founder of Butler Group which, prior to its acquisition, was Europe’s largest indigenous IT analyst firm. Business Intelligence, predictive analytics, big data, fast data, Enterprise Decision Management and all other technologies which aid business decision making are covered. info@butleranalytics.com