Streaming Analytics and Embedded BI
– The Route to Always On Smart Business Operations

By Mike Ferguson
Intelligent Business Strategies
February 2015

Prepared for: DATAWATCH
# Table of Contents

Introduction ........................................................................................................................................3
  What is ‘Always On’ Smart Business?.........................................................................................3
  Why Businesses Must Become Smart To Compete.................................................................4

Characteristics and Examples of a Smart Business .................................................................5
  Continuous Personalised Customer Engagement.................................................................5
  Continuous Optimisation of Business Operations ..............................................................6
  Continuous Risk Management............................................................................................6
  Continuous Dynamic Resource Planning and Financial Management............................7

Requirements – What’s Needed To Build An Always On Smart Business .........................8
  Multi-level Business Optimisation.......................................................................................8
  Data Connectivity...............................................................................................................8
  Streaming Analytics for Continuous Business Optimisation ...........................................8
  Combining Descriptive, Predictive and Prescriptive Analytics........................................10
  Embedded BI For Smarter Business Processes ............................................................10

Transitioning to A Smart Business Using Datawatch..........................................................11
  Datawatch Desktop...........................................................................................................11
  Datawatch Server............................................................................................................12

Conclusion ....................................................................................................................................13
INTRODUCTION

Most BI systems today typically provide insights about past activity

Businesses now want to go further with insights, predictions and recommendations provided in real-time

New data and new scalable platforms have emerged to produce deeper insights

Demand is now for insights to be available on a continuous basis increasingly in real time

In many organisations today, business intelligence (BI) is fairly mature with BI tools accessing data warehouses and data marts in support of both strategic and tactical decision making across many different departments. This is especially the case in marketing, sales and finance. Yet, despite the maturity of BI in the enterprise, most insight produced today is descriptive in nature in that it shows what has happened in the business over time. While this is useful, it is a long way off what most really want from analytical systems which is to go deeper by preparing and analysing any type of data, to deploy predictive¹ and prescriptive analytics² and to do more of it in real-time.

Not surprisingly, deeper insight requires new data, which means that new data sources of varying types are now in demand. This includes clickstream data, sensor data, text, and semi-structured data like JSON and XML. All of this is being captured or planned with data volumes growing rapidly and the rate at which data is being generated also skyrocketing. To cater for this, new big data platforms like NoSQL databases and Hadoop have entered the enterprise to help ingest and analyse more complex types of data at scale with much of the focus being on data science and complex analysis.

While all this is important, many companies also have an equally pressing need for insights to be available on a continuous basis to everyone in the employee base (and beyond) working in every day business processes to help them guide and optimise business operations. The vision today is therefore much more challenging. It is to become an ‘always on’ smart business.

WHAT IS ‘ALWAYS ON’ SMART BUSINESS?

So what exactly is an ‘always on’ smart business? This can be defined as

“Where business intelligence and analytics are used to guide people and applications so that they continuously know the best action to take and when to take it to dynamically (re-)optimise business operations and minimise risk in order to maximise profitability”

One of the key words in the above definition is the word continuously. To always know. To achieve this kind of vision means being able to derive insights from analysing any type of data, make new insights available anywhere, predict what might occur, and prescribe the right actions to take to keep the business optimised. But it is more than that. There is also a further requirement, which is to go beyond the current strategic and tactical decision making into continuously supporting the thousands of small decisions made in every day business operations. These small decisions should not be entirely reliant on business analysts to see everything. In the world of operations there is an increasing need to do much of this in near real-time. Add to that the need to cope with new high velocity data like sensor data and clickstream, and you quickly realise that achieving ‘smart’ operations in this kind of environment requires organisations to go beyond human analysis to:

• Automate the continuous monitoring of events as they occur in different parts of the business and at different levels in the business

¹ Predictive analytics is about forecasting what might happen in the future
² Prescriptive analytics suggest possible actions to take to reach desired outcomes
• Automate action taking where appropriate

• Embed BI into their core operational business processes to constantly alert and guide front line employees and applications to act in a more timely and effective way than they do today

The ultimate objective is to get to the point where analytical systems are continuously monitoring, managing and driving all business operations on a 24x365 basis.

**WHY BUSINESSES MUST BECOME SMART TO COMPETE**

Why do businesses need to become smart? The answer is simple – it is to survive, to drive new business opportunities, to continually engage customers, reduce risk, reduce costs and remain compliant. All of this is to increase competitive advantage. It is no longer just about business analysts, managers and executives making all decisions. It is about everyone pulling in the same direction including front-line operations staff. Even applications should participate in making decisions. In the case of applications, we mean allowing automated decisions to be taken when well-known patterns are detected.

In a way smart business is analogous to analysing a person’s health and fitness. The greater the investment in analytics to produce new and deep insight on a continuous basis, the more informed a business becomes and the more pro-active, responsive and competitive it can be when making decisions to improve its overall performance. This paper discusses how you can use streaming analytics and embedded BI to improve the overall fitness in your business.
CHARACTERISTICS AND EXAMPLES OF A SMART BUSINESS

There is no question that the characteristics of a smart business are different from that of a business with a data warehouse, and data marts that just uses descriptive BI. In a smart business there is a fundamental change in how BI and analytics are used that goes beyond strategic and tactical decision making to directly impact how a business operates. This includes tightening integration between operational and analytical systems. Usage of insight therefore goes beyond business analysts, managers and executives to directly include front-line operations staff, partners, suppliers and customers.

Data being analysed extends beyond processing historical structured transaction data stored in databases to also include real-time data-in-motion with both multi-structured and structured data now being embraced. Data ingestion has to scale to handle velocity i.e. the rate at which streaming data is being generated and arriving. With respect to analysis in the smart business, the increase in data velocity alone means that analyses cannot all be done by people. Guided and automated analyses are now part of every day business operations. Also actions should be taken by people AND by applications. The latter is possible because rules built-in to streaming analytical workflows can trigger automated actions. Examples include automated alerting, automated campaigns, automated re-ordering to prevent stock-outs in a retailer and automated recommendations.

In a smart business, insight is delivered via personalised role-based dashboards to browsers and mobile devices. This includes alerts, early warnings, real-time and historical trend visualisations, drill downs, guided intelligence, predictions and recommendations. On-demand access to insights and recommendations embedded in operational business processes and applications is also possible to guide people in their every day decision making.

To monitor business operations, ‘smart agents’ are deployed in different parts of the business on the look out for event correlations that, when detected, trigger event-driven automated analysis to analyse business impact assessment and automated action taking to prevent problems, reduce risk and to re-optimise operations. Time to action on alerts is also monitored so that automated escalation of alerts occurs if exceptions are not acted upon within a user-defined timeframe.

Furthermore, role-based scorecards with personal objectives, targets and KPIs can be organised into hierarchies at strategic, tactical and operational levels so that all KPIs roll-up to contribute to common strategic objectives and targets in a multi-level strategy management implementation. The purpose of this is to cause co-ordinated execution of a common business strategy across all levels of the enterprise.

There are so many examples of where streaming analytics makes sense in a smart business. They include fraud prevention (e.g. in financial services and in government), patient monitoring in healthcare intensive care and neonatal units, optimising logistics operations, monitoring and protecting against cyber security breaches, monitoring stock market trade activity, monitoring on-line gaming activity, monitoring on-line retail browsing activity and many more. A few examples are described below.

CONTINUOUS PERSONALISED CUSTOMER ENGAGEMENT

One of the most important areas for streaming analytics and embedded BI is in the front office. Organisations like retail banks, telecommunications and retailers are
finding that customers are increasingly interacting on-line as opposed to any other channel. In some cases, (e.g. on-line businesses), it is the only channel to interact with customers and so better on-line customer engagement is needed. Imagine if it was possible to offer personalised live recommendations to people on-line to improve customer engagement and help boost sales. To do this means being able to analyse live clickstream to monitor the on-line behaviour of prospects and customers as they browse. This is a truly real-time data and real-time analysis requirement. Not only does it require real-time but it also requires scale and total automation of data preparation and analysis if we are to monitor live web browsing from any device on a continuous basis, right around the clock.

Telecommunications companies can use this approach to monitor GPS sensors and live clickstream from smartphones in real-time in order to create new flash advertising services to retailers to trigger advertising based on a customer’s current location and/or their browsing behaviour.

**CONTINUOUS OPTIMISATION OF BUSINESS OPERATIONS**

One of the key uses of streaming analytics and embedded BI is in the ability to optimise and re-optimise business operations. Several industries have business operations that significantly benefit from continuous real-time monitoring to optimise efficiency and minimise cost. This includes retailer and manufacturing supply chain optimisation, manufacturing production line optimisation, distribution optimization, traffic flow optimisation in smart cities, utility grid utilisation and health monitoring. All of this is very much a new area where sensors are used to monitor operations. Also monitoring live trading activity in the financial and commodities markets is another example whereby competitive advantage often goes to those that can predict and respond the quickest. There are so many examples of business benefits of real-time analytics.

Going back to sensors, the emergence of sensor-driven smart meters has meant that utilities companies can now monitor grid energy consumption on a continuous basis. This allows them to more accurately match supply with demand in a much more timely manner and realise savings from not over supplying. Furthermore, real-time streaming analytics also allows utilities companies to introduce elastic pricing on a daily basis to increase revenue. This kind of smart business could be taken even further by using smart meters to improve customer retention and satisfaction. By using smart meters, utilities organisations could potentially help identify ‘energy hungry’ appliances in households and so make recommendations to help customers reduce energy consumption. This kind of pro-active service leads to increased customer satisfaction and customer loyalty.

**CONTINUOUS RISK MANAGEMENT**

In the same way that analysing streaming data-in-motion can be used to optimise business operations, it can also be use to prevent or minimise risk. For example to continuously prevent unplanned outages in machinery and other operating assets. A good example of risk management is in the Oil and Gas industry where the Deepwater Horizon oil spill\(^3\) (also known as the Macondo Blowout) in 2010 changed the industry. Today oil and gas companies have deployed sensors to monitor drilling as it happens and to also monitor well integrity on a continuous basis. All of this helps to minimise risk in terms of health and safety and also allows prediction of potential equipment failures long before they happen in order to prevent unplanned outages. This in turn helps optimise field service maintenance operations and to swing the balance toward preventative maintenance rather than trying to respond to unplanned failures after they happen.

---

\(^3\) [http://en.wikipedia.org/wiki/Deepwater_Horizon_oil_spill](http://en.wikipedia.org/wiki/Deepwater_Horizon_oil_spill)
Another example of continuous risk management is real-time fraud detection and prevention. Retail bank credit card fraud is a classic example. Also monitoring payments between parties together with unusual account activity all within a specific time window may indicate fraud.

**CONTINUOUS DYNAMIC RESOURCE PLANNING AND FINANCIAL MANAGEMENT**

Planning today is very often too slow, too expensive and not detailed enough. However streaming analytics can help to significantly improve this situation. By continuously monitoring business events in much more detail and on a real-time basis using statistical functions and forecasting, it becomes possible to understand key influences on driver metrics, continually re-calculate driver metrics to see and/or predict the impact on business performance KPIs. This in turn allows organisations to perform ‘what if’ scenario simulations more frequently to determine the best actions to take to keep a business on-track towards achieving its targets. It shortens planning cycles, facilitates dynamic adjustment of plans and allows for dynamic reallocation of resources.

Dynamic financial and demand planning become possible with dynamic resource allocation and more. This applies to both demand planning and financial planning as one impacts the other. With respect to demand planning, if a retailer monitors live sales at barcode scanning point-of-sale terminals in real-time, it becomes possible to see demand, predict stock outs, automatically re-order to match supply with demand, and monitor distribution to optimise just-in-time delivery. They could also dynamically manage supermarket human resources to perform shelf-restocking and point-of-sale activities. Of course there is a lot more to this than monitoring point-of-sale terminals. Data is also needed from systems controlling distribution allocation, inventory management, goods-in, claims management and returns.

Dynamic management of costs also becomes a possibility. With respect to financial planning, more dynamic control of costs can be managed using streaming analytics and embedded BI. One example is in the area of procurement where many companies have a need to better manage expenditure across their business. This is very challenging when managers in many different departments and geographies have spend authority. In this case, real-time analytics can be used to monitor personalised expenditure against budgets to prevent overspend and to avoid budgets being exceeded well ahead of plan.
REQUIREMENTS – WHAT’S NEEDED TO BUILD AN ALWAYS ON SMART BUSINESS

In order to create an ‘always on’ smart business, a foundational set of requirements need to be defined. These are discussed below:

MULTI-LEVEL BUSINESS OPTIMISATION

- The creation of a smart business requires co-ordination of insights, alerts, recommendations and actions by business users at strategic, tactical, and operational levels so that everyone associated with achieving a specific strategic objective has access to relevant information to help them contribute to common goals. That means that throughout the enterprise it should be possible to deliver dynamic role-based managerial scorecards and operational dashboards that combine relevant historical and real-time visualisations together with real-time alerts and recommendations to help people in different roles manage and operate the business more effectively. In addition, those tied to operational applications (e.g. bank clerks, contact centre staff etc.) should also be able to be guided by analytics and insights by embedding relevant analytics in applications and processes being used.

- Also, background automated stream processing should be capable of monitoring demand versus resources so that budgets and plans used at strategic, tactical and operational levels can be dynamically updated and/or changes recommended as part of a multi-level strategy management initiative. The purpose here is to ensure that everyone pulls in the same direction.

DATA CONNECTIVITY

- It should be possible to access multiple data sources for the purposes of analysis and decision-making. This includes structured historical data in traditional data warehouses, multi-structured data in big data platforms and also real-time data streams in stream processing servers.

- For the purposes of stream processing it should be possible to analyse data in motion⁴ as well as turn data at rest (e.g. data in a database) into data streams to add context to the analysis.

STREAMING ANALYTICS FOR CONTINUOUS BUSINESS OPTIMISATION

In order to define requirements for streaming analytics on data in motion, we first need to define what a streaming analytics platform is and what streaming analytics applications do. Figure 1 shows the basic steps in a stream processing application that would be deployed on a real-time stream processing server.

---

⁴ Wikipedia defines data-in-motion as data that is traversing a network or temporarily residing in computer memory to be read or updated. Data-in-motion is therefore data that has been created but not yet stored.
Stream processing analytics can automatically analyse data from one or more data streams and make automated decisions.

They can also make data available to third party tools for visualisation.

Data preparation is fundamental to streaming analytics.

Automated analysis is needed to cope with high velocity data.

Scalability is also needed.

Automated decisions should also be an option.

Alerting is a fundamental requirement.

Access to streaming data by visual discovery tools is also needed.

Figure 1

One or more data streams can be accessed and data prepared for analysis. Continuous queries and analytics are then applied to the data to produce a desired result. Note that analysis can be automated by embedding analytics in the stream processing application. Alternatively, third party tools could query the streaming data for visual discovery. On this basis, the requirements for streaming analytics can be defined as:

- It should be possible to prepare data from one or more data streams for the purposes of analysis. This includes filtering, cleaning, joining and compressing streaming data to produce the set of variables needed for automated analysis.

- It should be possible to automatically analyse streaming data-in-motion looking for correlations as soon as possible after events occur to monitor operational activity as it happens i.e. event-driven automated analysis is needed.

- It should be possible to use predictive and statistical models to support automatic analysis of data-in-motion during stream processing.

- It should be possible to parallelise the execution of predictive and statistical models in streaming analytic applications to scale to handle very high velocity data.

- It should be possible to analyse structured and multi-structured data during real-time stream processing.

- It should be possible to update predictive and statistical models on-demand or at selective intervals.

- It should be possible to support rule-driven automatic decision making via a rules engine accessing the outcomes of predictive / statistical models during real-time or near real-time stream processing.

- It should be possible to automatically invoke alerting services, transaction services and/or whole business process workflow services as part of an automated action during real-time stream processing.

- It should be possible for visual discovery tools to connect to high velocity stream processing platforms and to produce new insights such as aggregates, high-level statistics, or adjustments to key operating metrics.
• It should also be possible for visual discovery servers to receive alerts, automatic recommendations and live data pushed to them by stream processing servers

**COMBINING DESCRIPTIVE, PREDICTIVE AND PRESCRIPTIVE ANALYTICS**

- It should be possible to combine role-based descriptive, predictive and prescriptive analytics on the same managerial or operational dashboards. This includes integration of real-time event data and alerts into visualisations to provide contextual early warning alerts etc., alongside historical data. This capability allows single place for managers and front-line operations staff to get all the information they need to drive decisions

- It should be possible to automatically combine historical and real-time analysis to assess the business impact to help guide users to making the right decisions

- If possible, organisation structure needs to be understood to escalate alerts if an action has not been taken by a user within a user-defined timeframe

**EMBEDDED BI FOR SMARTER BUSINESS PROCESSES**

- It should be possible to publish analytical artifacts such as queries, reports, dashboard components, predictive models, etc., as RESTful services that can be embedded in applications and processes for invocation on-demand. Embedded BI can be achieved via integration of BI servers with Enterprise Service Bus (ESB), business process management (BPM) and other infrastructure technology to help make insights more pervasive

- It should be possible for users in different parts of the enterprise to invoke and access published analytical artifacts on-demand to help them perform tasks more effectively and guide every day actions

- It should be possible to invoke in-database, in-hadoop and in-stream based predictive and statistical models from within operational applications and business processes to automatically analyse historical, multi-structured and real-time streaming data on-demand

- It should be possible to support on-demand recommendations in business processes i.e. online requests for automated decisions to guide people in operations

- If operational applications requesting on-demand analytical services are available 24x365 then BI servers and underlying data must meet the same availability requirement

- It should be possible to support large numbers of concurrent users invoking analytical services on-demand from desktop, web and mobile device based operational applications

- It should be capable of supporting large numbers of concurrent requests from applications for business insights that depend on automated analysis

- It should be possible to combine automated analysis and rule-driven automated actions to create *decision services* that can be invoked on-demand, on an event driven basis. An example of this is a recommendation. Recommendation services are a key component in any smart business

- In addition to on-demand and event-driven decision services, it should be possible to schedule automated analysis and action taking at user defined intervals to automatically identify opportunities on a timer-driven basis
TRANSITIONING TO A SMART BUSINESS USING DATAWATCH

Having looked at streaming analytics and embedded BI, this section of the paper looks at one vendor’s technology to see how it rises to meet these requirements and deliver real-time streaming and embedded actionable insights to all parts of the enterprise. That vendor is Datawatch.

Datawatch was founded in 1985 and provides visual discovery and dashboard software. Datawatch products include:

- Datawatch Desktop
- Datawatch Server
- Datawatch Report Mining Server

The focus of this paper is the Datawatch Desktop and Datawatch Server. Datawatch Report Mining Server will be covered in another paper.

DATAWATCH DESKTOP

Datawatch Desktop is a visual discovery and dashboard building tool that connects to personal data stores, database, applications and real-time streaming data sources including:

- Microsoft Excel
- Text files
- OData
- Big Data NoSQL databases like Cassandra, MongoDB and IBM Cloudant
- Big Data search engine technologies like Splunk
- JSON data
- XML data
- Salesforce.com
- OSISoft PI System - used extensively in industries like Oil and Gas, Petrochemicals, Power and Utilities and other verticals
- Relational databases
- Multi-dimensional databases including IBM Cognos TM1 and others (via MDX)
- Complex Event Processing (CEP) stream processing platforms such as Tibco Streambase, SAP Sybase ESP, IBM InfoSphere Streams

Also connections to Apache Kafka/Storm, Spark Streaming, Microsoft Azure and Google Analytics are all in development.

Looking at the above mix of data sources, business analysts using Datawatch Desktop can create workbooks that analyse historical data at strategic and tactical levels as well as creating workbook visualisations of real-time data at the operational level. They can then combine these visualisations onto the same dashboard. This means role-based dashboards can be produced that concurrently monitor strategic, tactical and real-time operational data. Potential information consumers, such as managers, can therefore see what is going on in
Historical and real-time visualisations can appear on the same dashboard

Role-based dashboards can be created that are aligned with business strategy

Time-series visualisations are available to monitor real-time activity within specific time windows

Custom analytics developed in R and Python can also be invoked

Streaming Analytics and Embedded BI

their area of responsibility right now alongside trends at a tactical departmental or business unit level as well as the overall business impact at a strategic key performance indicator (KPI) level. This kind of capability allows strategic, tactical and operational insights to be associated with specific objectives in a business strategy to be combined. For example, in a business strategy the board of a business will likely define strategic business objectives and targets. These might be categorised for easier understanding by customer, finance, risk and operations. By creating role-based dashboards containing strategic, tactical and operational insights for use by front-office customer oriented managers, financial managers, risk managers and operations managers available it becomes possible to align actionable insights with business strategy objectives and targets. In this way, people at all levels have everything available at a glance and all can work together towards common goals.

Datawatch Desktop has the capability to do this. It connects to historical data and real-time data via third party streaming analytic applications that automatically analyse data as it arrives (as per Figure 1) and/or offer up data-in-motion for query and analysis in Datawatch Desktop. In this capacity, Datawatch Desktop is querying real-time data streams and receiving the results of such analyses for visualisation. Datawatch therefore offers business analysts the ability to do visual discovery on real-time and historical data as well as to create dashboards for distribution to information consumers. Given that real-time streaming data is frequently time-series oriented, Datawatch Desktop provides a set of time-series visualisations to facilitate time-series analysis of the streaming data. The tool is also extensible in that Datawatch supports direct connectivity to analytics platforms such as RServer and Python server where custom built predictive analytics and statistical models can be applied to data once it has been retrieved from a data source. This is true both in Datawatch Desktop and Datawatch Server.

**DATAWATCH SERVER**

Dashboards can be shared via Datawatch Server

Data can be held in-memory to support large numbers of concurrent users doing interactive analysis

Support for embedded BI and alerting help people working in front-line operations become more effective

Once business analysts are happy with the dashboards they have created, they can share them with others by publishing these dashboards to the Datawatch Server for information consumers to access via browser or mobile device.

In order to support concurrent users and improve dashboard interactivity, Datawatch Server also includes an in-memory database that can be used to cache data when dealing with slow data sources.

It is also possible to embed BI published to the Datawatch Server in other applications. This is because Datawatch Server supports a REST API that can be used to control the dashboards and to integrate them into other applications.

In order to help guide and optimise business operations Datawatch can generate real-time alerts that can be set up to visually notify the user on the dashboard. Alerts can also be pushed out to users via email. This kind of support helps organisations manage by exception rather than try to analyse everything that has happened.

In addition, given that Datawatch can also analyse documents, it is also possible to store and manage those documents and corresponding extraction models on the Datawatch Server.
CONCLUSION

Organisations are now looking to work smarter in all areas of their business. They also want to do this at all levels from executives at strategic levels to front-line workers in operations. To make that happen, analytical systems need to be integrated into every business process so that insights can be made available in the context of each specific task being performed. In that sense, analytical systems are being surrounded by more and more applications that need to access them. This is shown in Figure 2.

To compete, organisations have to monitor their operations and integrate insights into their operational business processes, managerial and operational dashboards to make sure that people are always aware of what is happening in the business at all times to be able to make informed decisions.

People want the right insights to be available in the context of every task. Therefore embedded BI is a must in any smart business. In addition, streaming analytics is also a must. This is needed to automatically monitor events to quickly spot opportunities, avoid risks and re-optimise the business operations as and when needed. Real-time visualisation is also critical to confirm what is happening now and see patterns in data. It is particularly relevant in operations. Finally, role-based dashboards that combine real-time and historical data need to be published and made available at all levels in the enterprise.

Datawatch is providing all of these things in its visual discover tools that tap into historical and real-time data feeds. It supports role-based dashboards that combine historical and real-time data. In addition, it supports APIs on Datawatch Server to embed insights in applications and processes. Finally, it is expanding its products to support new real-time industrial, financial and big data streaming analytics platforms that allow it to move into new markets like Internet of Things. All of this is fundamental to becoming a smart business and makes Datawatch a candidate for any organisation seeking to make that transition.

Figure 2
About Intelligent Business Strategies

Intelligent Business Strategies is a research and consulting company whose goal is to help companies understand and exploit new developments in business intelligence, analytical processing and enterprise business integration. Together, these technologies help an organisation become an intelligent business.

Author

Mike Ferguson is Managing Director of Intelligent Business Strategies Limited. As an independent analyst and consultant he specialises in Big Data, business intelligence, analytics and data management. With over 33 years of IT experience, Mike has consulted for dozens of companies on business intelligence, enterprise architecture, business integration and data management. He has spoken at events all over the world and written numerous articles. Formerly he was a principal and co-founder of Codd and Date Europe Limited – the inventors of the Relational Model, a Chief Architect at Teradata on the Teradata DBMS and European Managing Director of Database Associates. He teaches popular master classes in Big Data Analytics, Self-Service Business Intelligence, Data Virtualisation, Enterprise Information Management and Master Data Management.