Business Intelligence SaaS
Pros and Cons

A DMG Federal eBook

Keith Boyer
March 2013
New service models have appeared as virtualization has gathered steam and cloud-based systems have become a dominant paradigm in the world of information technology. Cloud-based service models are often denoted “Anything as a Service” or “XaaS” (where X represents whatever is being offered). The three most common examples are Software (SaaS), Platform (PaaS), and Infrastructure (IaaS). This group (software, platform, and infrastructure as a service) is often itself shortened to the acronym SPI.

SaaS is becoming increasingly popular, especially in the context of business intelligence (BI); the science of analyzing data in order to understand the factors influencing an enterprise's fortunes better. BI can involve looking at general trends in a company's field of operation, consumer attitudes to its brands, and/or price changes affecting the raw materials used in its products.

It could be stated reasonably that the last few years have brought a crisis for BI. Part of the reason for this has been the phenomenon of Big Data; the name given to data sets that cannot be handled by conventional database procedures. These data sets are too vast, too rapidly-changing, too varied, or a combination of these factors. The escalating quantities and changing types of data have been a side effect of the massive upheavals that Information Technology (IT) has seen in the last 15 years. Of course, the global ubiquity of the Internet and the widespread adoption of cloud-based computing strategies play a great role in the influx of big data.

Where Does Big Data Come From?
This flood of data is fed by a multitude of sources; therefore, to simplify the picture a little, these sources can be grouped under three broad categories.

1. Internally Generated Data
Internally-generated data frequently originates with activity in the IT systems of a company. It can involve applications logs, information about network traffic, and what has come to be known as “clickstream data.” This is information regarding customer (or potential customer) activity online. Historically, it was really only possible to harvest simple chunks of data relating to an individual customer, such as what they bought and how much they spent, with possibly a little extra detail, such as what sector the customer operated in. Now the amount of data available for any single individual, even from an isolated transaction, is greater than ever. The availability of large quantities of low-level, internally-generated data for capture and analysis has the potential to revolutionize the way enterprises operate, let alone how they view themselves and their clients.
Clickstream data is a good example of the ways in which data has mutated and multiplied in recent years. It includes data on how a customer got to a company's website, what they were looking for, what they looked at while they were on the site, and any purchases they might have made. From this data, it's possible to deduce important facts about how customers use a site, whether or not it's providing them with what they need, and how it can be improved. In addition, this information can generate new marketing ideas for product or service placement.

2. Data from Data-set Marketplaces
Data-set marketplaces provide a clearing-house for data sets acquired from various providers. One example is Infochimps, the company behind the Infochimps Data Marketplace. This marketplace allows clients access to tens of thousands of data sets drawn from hundreds of thousands of different sources. Data sources used by Infochimps Data Marketplace include Wikipedia, Foursquare check-ins, answers to OK Cupid questionnaires, and information from various national databases.

Like other data marketplaces, Infochimps collects a lot of its data sets from social media. In fact, the company breaks down its data sets into two general categories; each served by a particular range of advanced programming interfaces (APIs): Geo data and Social data. Geo data empowers companies to analyze data sets as they relate to specific geographical locations, enabling the creation of hyper-local content, for example. Social data can be used to monitor the health of a brand in the social media sphere, to spot spammers and other abusers of social media platforms, or to gain a better understanding of the needs and wants of a particular segment of the customer base.

3. Third-party Data Generators
Third-party data generators are entities that specialize in the collection and provision of data to various clients. They include international bodies such as the European Union, governments both national (including the U.S. Federal Government) and regional, such as states or counties. Such bodies generate enormous quantities of data on everything from public health to commerce to education.

Third-party generators can also include commercial operations that harvest data for private clients. For example, many companies’ main activity is compiling, distributing, and processing of consumer surveys. These, along with companies involved in opinion polling, can be a rich source of important data. An illustration of a third-party generator is Dun and Bradstreet, who collects and distributes business credit information and credit reports.

Aside from these three categories, other data streams may include social networks and other social media, collected by the company itself rather than a data-set marketplace.

Coping With the Data Revolution
All this data is both a blessing and a curse to modern enterprise: it is potentially of immense value but sprawling, frequently unstructured, and often so unwieldy as to be impervious to conventional database methods or analysis. The volume of data involved is so enormous that attempts to tackle it using conventional BI initiatives can take upwards of a year to complete. This grindingly slow pace is less and less acceptable in a business world where responsiveness is ever-more important.
In the 2010s, simply eschewing big data as a source of intelligence could hardly be described as a solution. It should be observed that traditional intelligence initiatives tend to be long-term projects that may not yield results for months. Even if that wasn’t the case, walking away from the goldmine of actionable information offered by big data could be economic suicide. A whole industry sector has sprung up around big data as it relates to BI, with established companies and feisty startups vying to offer the most effective solutions for the challenges posed by the modern data landscape. This indicates a significant need for such solutions, and a considerable advantage for those who can deploy them effectively.

Novel approaches have emerged that are specifically tailored to the demands of new informational phenomena. Software implementations of these approaches can produce effective applications for generating intelligence; however, such software is becoming increasingly demanding, as the amount and complexity of the data that must be handled keeps growing. Instead of attempting to support these resource-hungry software applications in-house, more and more businesses are turning to cloud BI.

Advantages
In this model, the installation and maintenance of the software is undertaken by a specialist service provider. Companies can then outsource to this provider on whatever basis is most suitable. They may set up a long-term arrangement wherein they have continuous access to an instance of the application that's been fully configured to their specific needs; alternatively, they may opt for a pay-as-you-go or pay-per-use arrangement. Like other cloud computing services, this outsourced model has the following advantages:

- Being freely scalable.
- Obviating the need to invest in specialized software or infrastructure.
- Eliminating the necessity of hiring specialist staff to run the installation.

A company offering analysis through the SaaS model can focus on developing their software applications to the fullest and can invest in high-end infrastructure, often to a degree that's impractical for most other businesses. It’s much easier for a SaaS-based company to attract support in adopting the latest new technology than it is for an enterprise whose focus is elsewhere to justify such an outlay to their shareholders.

Talent and Skills
Additionally, a company specializing in SaaS BI may have the advantage over an IT department in a larger company when it comes to staff. In theory at least, an XaaS BI-centric enterprise is best placed to know which skillsets are going to be the most valuable and to spot those individuals with the greatest potential. It can also offer the kind of unique opportunities and challenges that may attract the top talent in the field.
Given the potential advantages of SPI-based intelligence in general and SaaS in particular, it’s no wonder that solutions based on this model have grown in popularity over the last few years. As far back as 2009, the Aberdeen Group (a company specializing in fact-based research and predictive analytics) noted that the uptake of business analytics via SaaS had more than doubled over the previous two years. That trend has continued, driven to a large degree by smaller companies. Small and midsize businesses have eagerly adopted analytics in general; and have chosen analytics provided via SPI as the most effective means of obtaining the information that they need.

It’s not hard to understand why. The smaller a company, the more it needs the advantages that can accrue from business analytics; yet the less it can afford to invest in on-site infrastructure, specialist software applications, and new staff - particularly new staff with what is still something of an esoteric skillset. These and other factors make the kind of one-off or scaled usage models offered by cloud BI tremendously attractive to Subject Matter Experts (SMEs). As the Aberdeen Group notes, the result has been a kind of viral distribution of the SaaS BI meme among enterprises of all shapes and sizes, from local startups to multinationals.

**Actionable Information**
A question that many will ask is, “What is all this data good for?” Big data can be converted into actionable information in a number of ways, from developing predictions about future trends to data visualization. These are just two examples of how new approaches to data can help a company meet its key performance indicators (KPIs).

A crucial application of big-data-derived BI is in the field of predictive analytics: the science of making predictions of future events based on past behavior and events. As more and more sophisticated analytical tools emerge, the number of data sources that can be factored into an analytical operation keeps growing. In addition, relatively novel data sources that are now available, such as social media, can be incorporated into analytical operations. This is anticipated to yield far more accurate predictions about future trends, whether it’s the performance of a particular commodity on the global financial stage or the behavior of a particular customer base.

**Key Performance Indicators**
By analyzing big data sources, a company can better track its fulfillment of its (KPIs). Better yet, new analytic tools can help an enterprise to determine more accurate and appropriate indicators of performance, offering a more realistic and useful set of metrics by which a company can track its ongoing effectiveness and development.

**Visualizing Data**
Another key role for big data analytics is in providing data visualization. Streams of figures and abstract data are fine for computer systems; they’re not so good for human beings. Clear, well-presented graphical data is increasingly important whether you are trying to get a clearer understanding of complex data sets yourself, presenting information to help sway potential investors, or creating interdepartmental reports that will be used to make business decisions. Almost any business can benefit
from analytics that offer visualization; with sophisticated cloud-based software now available on a pay-as-you-go basis, this kind of functionality is now within the reach of even the smallest enterprises.

An example of this type of offering is Tableau Software, a company specializing in data visualization. Solutions integrating Tableau's software with other analytical applications are available in a variety of options: pay-per-use, pay-as-you go, and standalone installation basis.

**Business Case: RS Components and SAS**

The field of e-commerce, particularly those enterprises that rely to some degree on demand shaping, is full of examples. One of the most notable of these has been RS Components, an electronics vendor that has made extensive use of new business analytics in demand management. A major part of the company's business model involves persuading customers to purchase more plentiful alternatives when stocks of their desired option are running low, thus ensuring a greater number of satisfied customers.

Originally, the work of analyzing demand and determining the needs of the company's customers was performed with traditional methods: data was handed off to analysts, who would deliver intelligence that RS could turn into initiatives such as price changes, promotional activity, and so on. Given the masses of data generated by their e-commerce platform and the limitations of traditional analytics, this became unwieldy. The analysis itself was consuming vast amounts of time, making the company less and less responsive and leaving insufficient time to develop and promote new initiatives. By adopting innovative SaaS analysis methods, the company was able to boost repeat visits to its e-commerce platform by a full 20 percent.

**Red Flags: the Downsides of SaaS BI**

We've established that there are a great many advantages to utilizing SaaS BI. Like any technology, however, cloud-based BI solutions do have their downsides. It's important to consider potential problems or contraindications before adopting a solution.

**Security and Regulation**

Any cloud-related outsourcing operation requires careful consideration of any possible security ramifications. This is very much the case with business analytics operations, which may involve sensitive information relating to the enterprise running the analysis, or possibly clients' private data. This last should be of particular concern; the handling of personal information is subject to various laws and regulations, potentially leaving a company open to regulatory or legal sanctions in the event of a data breach. Moreover, if it comes to light that a company has been lax in the matter of protecting customers' data, the resulting loss of trust can be catastrophic.

In heavily regulated industries such as financial services or health care, it may be all but impossible to make use of SaaS BI without violating regulations. In many cases it's hard to avoid a situation
where data effectively passes into the hands of an unauthorized party. Unless the vendor can absolutely guarantee regulatory compliance at all levels and at every stage of the project, it may be necessary to eschew such outsourcing and look at alternatives, such as implementing an autogenous solution on a private cloud.

Whether or not there's a regulatory issue to scrutinize, security should always be a top consideration when using any cloud-based service. Even if a company isn't using customer records as a data source, there's still a risk that sensitive data could be compromised if a service provider isn't sufficiently security conscious.

**Cost of the Cloud**

Another potential issue is cost. While it's typical for cloud-based analytical solutions to be significantly less expensive than in-house solutions or outsourced solutions using traditional methodologies, this is not the case in every instance. Companies looking to save money by using cloud-based software applications in place of traditional intelligence analysis or in place of autogenous analytical software installations need to be aware that “cloud” doesn't always translate into “cheap.”

Vendor prices are not always transparent, with hidden costs lurking in the small print to trap the unwary. Data overages, additional premium services, and other extras can all push the price of a project well over and above what the company was expecting to spend. The overall cost of a particular solution will depend on a number of variables, one of which is the number of data sources used. The more sources are involved, the more a company will end up paying for a particular analytical operation. The data sets themselves may be expensive to obtain; more to the point, however, is the fact that more data sources translates into more processing resources when it comes to conducting the required analysis. Since cloud costs scale with resources used, every extra source can potentially push up the price that a company pays.

Even if a cloud-based solution turns out to be more costly than a similar operation conducted using traditional methods, the extra expense may still be justified if it produces a faster result or more detailed and accurate intelligence. Before adopting a cloud-based solution, it's vital that the business conducts a thorough cost breakdown, weighing up the expected gains against a detailed list of the actual costs of the project.

**Leading the Field: Analysts, Vendors and Platform Providers**

We've already touched on the work of the Aberdeen Group in developing accurate predictions of the way in which cloud-based intelligence would come to dominate the field of business analytics. Let's take a look at some of the other big figures in the big data landscape:

Like the Aberdeen Group, Gartner specializes in providing fact-based analysis. They differ from Aberdeen in that their focus is specifically on the field of IT. Gartner offers consulting services to thousands of enterprises; in addition, they also publish original research based on their analytical work.
Business Intelligence SaaS Pros and Cons

One company that's currently thriving at the conjunction of big data and cloud computing is Kognitio. Founded in the late 80s, Kognitio has always specialized in the provision of infrastructure and platforms for data analysis. They now bring this experience to bear in providing a platform for modern analytical software, attracting clients from many different areas of industry. A cornerstone of Kognitio's appeal is its adoption of in-memory processing, which helps deliver much faster results than conventional disk-based infrastructure. They offer their platform as both a stand-alone installation that a company can run on its private cloud and as a scalable pay-as-you-go service on a public cloud.

Another vendor offering platform and infrastructure services, ParAccel offers analytic databases, data warehouse resources, and cloud analytics. ParAccel is noted for providing analytics that are tailored to handle Hadoop data.

Birst is an award-winning SaaS vendor. It specializes in providing business analytical solutions. Birst describes their approach as “agile analysis:” their software can respond to and incorporate the kind of cross-platform, multi-departmental, and inter-deployment data sources with which traditional analytic methods struggle. Some of Birst's high-profile customers include Rackspace, Citrix, and EMT associates, a provider of specialist emergency training.

Conclusion
The benefits of SaaS BI include lower costs in many instances; scalability; wide availability; and a broad range of services and solutions. The potential downsides include security and regulatory concerns as well as the possibility of hidden costs. Before you adopt any solution, it's vital to weigh the various advantages and disadvantages with your business needs.

References
http://searchcloudcomputing.techtarget.com/definition/Software-as-a-Service
http://www.kognitio.com/assets/images/diagrams/cloud_timeline_c.jpg
http://www.yellowfinbi.com/YFCommunityNews-Aberdeen-Business-Intelligence-technologies-to-handle-Big-Data-107097
https://www.sas.com/success/rscomponents.htm