

# An Intelligent way to do Business

Tom Alford asks: 'Is BI just another acronym to add to the list, or is it a valuable tool to make sense and further profit from your data?'

The problem with data is that there is too much of it. Most organisations generate and store vast amounts for their day to day operations. Gartner research from 2002 has suggested that, on average, companies manage 14 databases and spend between 60 to 70 per cent of application development budget trying to find ways to link and gain access to these databases.

Storing data is not the problem. Even within simple organisational structures, data sharing is the real issue. Different business functions have their own set of key metrics or 'truths' about each product, so duplicated information, conflicting results and confused decision-making thrive.

Hard-wiring data sources into one centralised area is a start. But what happens when you want to change it? And even with neatly interfaced silos, the data is useless unless it can be systematically collated, presented and acted upon quickly. As Richard Neale, product marketing manager for Business Objects, says: 'It's like having a warehouse full of food and letting 80 per cent of it go rotten.' Vital questions such as: 'Are we producing products for regions in which demand has fallen?', 'How much return can we expect from marketing to a particular segment?', or even more urgent questions relating to exposure and risk, need answering now, not later.

The answer to many data analytics prayers seems to lie in Business Intelligence (BI) platforms. BI as a phrase was coined and popularised by Gartner. Richard Peddar, CEO of Granville Associates, a company of ex-bankers offering an OLAP-based BI solution, defines it as: 'The technique allowing management to get at their own data, understand it and use it to help make better decisions and ultimately keep their customers happy and make more money'. This seems to be the consensus among vendors. Users, too, seem to share this aim. As Achim Vogt, COO of Deutsche Postbank, London, says of his operation's foray into BI: 'We want data consistency, and we want to save

money.'

Certainly for vendors (there are about a dozen significant players in the field) it seems to be a satisfactory definition. Although in practice it is a fairly vague generic term, covering a wide range of applications and systems, the market for BI systems is currently estimated by Gartner to be worth around \$600 million a year in Europe alone. For the end-user, BI needs closer examination to see if it really is the intelligent option.

## A poor track record

BI software, in its various forms, has been around since the early 90s. A lot of BI technology which was developed in those early days failed to match the promise. Expensive tools that often didn't deliver, or only partly lived up to the hype, damaged the image of the industry. The residual mistrust is understandable. Richard Green, COO of Co-operative Bank Treasury, comments on a previous encounter: 'We often had to call in Oracle experts to access and manipulate our data. I found even fairly intelligent users couldn't get at it.' The biggest risk, he feels, was staff cover. Where the expertise was so narrow, a bank became highly dependent on just one or two individuals.

A level of suspicion developed over time, with criticisms levelled at the concept. As BI software purchase often occurred at departmental level, many companies had a number of conflicting systems in operation. Even as late as 2002, a report from Forrester Research pointed to 80 per cent of its survey sample using multiple BI applications. Buying in tools on a per-user licence sometimes meant licences were left to gather dust where companies had over-estimated the number of likely users. Lack of productivity pushed return on investment even further away.

Another pitfall for multiple BI packages within one company was the increased demand for training, development and support services. Although vendors used

the same language to describe functionality, there the similarity ended. The tools themselves varied considerably. Different BI software and different data sources were used to produce department-specific results. Some data fields crossed the boundaries between departments, for revenue and profit, for example. But without a coherent analytical package, it was possible for conflicting results to be produced even from the same data. In a 2003 Forrester Research report, 53 per cent of senior operational and financial managers cited fragmented or inconsistent data as a major problem.

## What has changed?

So, not a very compelling case for BI tools you may think. The problem lies in companies embarking on BI projects without any recourse to planning. Now though, a number of vendors are offering broad BI platforms. A platform pieces together components and crucially offers a unified data layer, drawn from the client's disparate operational systems. This layer takes the form of metadata - data about data - stored in a virtual warehouse. Exactly the same key data can be accessed through a browser, LAN or WAN, by any authorised personnel within an organisation, using a variety of BI tools to manipulate and present it.

With shared metrics, courtesy of the platform, users are able to discuss the same key data and finally abolish the arguments over whose tools are best and whose data is more accurate. Advocating the positive effects of BI on Coop's treasury front-to-back office operation, Green says: 'To be able to take extracts, control and eliminate duplication, write reconciliation reports off them, put them into one database and then apply quick access tools to get reasonable business information out is very valuable.' Achim Vogt at Deutsche Postbank agrees: 'At first we were just curious, but we quickly realised its potential. It is a very flexible tool.'

But if BI platform vendors really are offering a 'complete solution', why hasn't uptake been more voracious? 'It's still a cultural issue for some,' says Neale. 'BI is often mistakenly seen as a technical job, not a business job. It is seen as a tool for power users and analysts only.'

Peddar often sees a similar suspicion of BI. 'Deep cynicism is one of the first reactions from banks.' This, he says, is perhaps based on the banks' lack of current knowledge of BI platforms. It may also be that some staff, particularly in middle management ranks, offer resistance, believing this technology could undermine their roles. There is also a perception that integrating a system involves too much risk and cost.

There is no reason for reticence now, argues Peddar. The technology that did work has been developed in the real world. It has really come of age. He firmly believes there is a new wave of interest in BI. As an experienced user, Green concurs: 'It seems very real to me that we can now empower users.'

#### Application areas in banks

A reason to explore the option once more may have been found. In the financial sector, there is an ever increasing need for 'business agility'. A powerful lure should be to chart a path through turbulent economic waters by rapidly delivering more accurate metrics. Jofi Alexander, marketing director of BI vendor, In Group, says: 'Senior management need to make sure they are achieving their strategic goal. If they aren't, they need to know why.' It is important to find out if assumptions made in the creation of that strategic direction were correct.

Senior executives and branch managers at West Bromwich Building Society use Business Objects' Dashboard Manager BI system. 'This gives a whole snapshot of company progress,' says systems development manager, David Eggleston. 'Executives can easily drill down to where the problem is.' For example, an alert is flagged where internal limits are in danger of being exceeded, such as where commercial mortgage offerings are about to exceed the amount West Bromwich can offer as a percentage of its residential mortgage business.



Jofi Alexander,  
In Group

Most BI systems will enable the user to focus on data for a specific product, area, branch, or even sales person. This makes it easier and quicker to react to a situation, maybe changing or withdrawing a particular product. 'In the past it could have taken a month to get a huge paper-based report,' says Eggleston. 'By this time it could be too late.'

Perhaps an even more urgent driver is increasingly stringent corporate governance. 'Managers are asking questions about whether they can trust the data they are using,' says Neale. 'It is becoming a boardroom issue now.' Sarbanes Oxley, for example, holds individual board members responsible for corporate offences under the Act. Huge financial penalties and custodial sentences have already been handed down to guilty parties. Alexander cites the Enron scandal as a classic example. There is an increasingly watchful and suspicious eye on corporate activity. 'Members of the board are now really putting their names on the line by saying, I believe in these numbers. But they're looking at numbers that have potentially gone through several iterations of interference. Without BI it is very hard to go back and ask where those numbers have come from.'

Boards now have to sign off what is essentially an internal control questionnaire, which is auditable. Although Sarbanes Oxley only applies to institutions in the US, or those dealing with the US, Peddar believes versions of this Act will soon be on

the statute books in Europe. 'This worries people. Often there is no tied or disciplined connection between data at the operational level and the information the board is looking at.' If someone gets a feed out of a controlled system and populates an uncontrolled Excel workbook and uses that data, the audit trail is lost.

BI platforms consolidate and reconcile data. In doing so a full audit trail of any data change or movement is compiled, down to a granular level. Although it is unlikely a system will automatically map data to the required format for external reporting, the data delivered is in theory acceptable for the task (see box, p35).

#### The components of BI

So what else is the integrated BI platform offering? To understand this it is necessary to take a closer look at the technology involved, using the spreadsheet viewpoint as an example.

Traditionally, once extracted from the core operational system, a bank's financial data such as budgets, forecasting, and P/L accounts, is almost universally worked on in a multitude of spreadsheets (most likely Excel). The problem with multiple spreadsheets is that presentation can become rather messy when multiple dimensions are opened up for comparison. For example, an accountant may use data to create budgets, measure actuals, and then re-forecast. One workbook may be P/L measured against time. Another workbook may be budgets measured against actuals and forecasts. The user may then need to drill down into the data, accessing sales worksheets, or products and pricing, and then relate this back to the higher level data. In doing this the user may quickly run up against limitations of the spreadsheet, ending up with multiple workbooks that somehow need tying together.

Obtaining fundamental data, coming from perhaps a sales system, a product database, a customer database, and last year's financial transactions, then sucking it all into one place and manipulating it, becomes taxing. A spreadsheet will be able to do this if links are created. But these links are really just one-offs. The user has to remember to refresh the data every time the underlying system changes. You end up with a rats' nest, or what some people call



Richard Neale,  
Business Objects

**'It's like having a warehouse full of food and letting 80 per cent of it go rotten.'**

'Excel hell', says Peddar.

Rigorous accounting procedures may overcome the problem at this level, but moving away into more operational procedures starts to reveal the weaknesses of the process. 'One of the common problems in banks is that everybody is sucking data from different systems, populating their workbooks, massaging data and reaching a conclusion,' says Peddar. This information can then be passed on to, and be manipulated by, several other people. In the end, he says, 'You've got data coming up from all directions, all with different analysis based on each departmental perspective. The end result is multiple versions of the truth.'

BI in part aims to tidy up the delivery. But although many of the applications can run directly against operational systems, these are over-complex and relatively slow in the context of BI. The perceived wisdom today is to start by building a virtual data warehouse. The warehouse is fed by the underlying data in the operational system. However, it acts as an intermediary storage depot, populated with simple data structures that can be browsed easily by non-technical, 'casual' users. Data warehouses always exist separately from the operational system simply because they have a fundamentally different architecture and purpose.

#### **Extract, Transform and Load**

To move the data from the operational system into the warehouse, data integration software is needed. The most common tool is an Extract, Transform and Load (ETL) application. This must work to a pre-configured data model. As this forms the basis of a bank's data source for BI, it must be executed with care. In a polite version of the old adage, Alexander says: 'The quality of the ETL process determines the quality of the data in the system.'

Most vendors, or their consultants, will assist in identifying sources of key data to populate the warehouse. In the past this job may have been considered from a technical perspective. This was unlikely to take into consideration end-user needs. The task needs to be driven by business demands to succeed. A bank must understand what it is doing with the data, and what it would like to do. 'We like to stretch our clients' thinking on this,' says Alexander.

At this stage it is important not to underestimate the benefit of a vendor with experience in the banking field. It will understand what is important about this specialised field and should be able to offer constructive advice on building what is effectively a blueprint for the BI system.

Why so important? Typically a bank will regularly only look at a small part of its data haul - consider that something like

100,000 data fields in a proprietary core system is not unheard of, and you realise why. Part of what the software does is examine individual data fields and apply rules to consistently convert the contents to the form required by the data warehouse.

The problem of finding the right data is amplified because, although a Universal Naming Convention now exists, a prior lack of standardisation across the industry has led to fields or paths possibly having different names, depending on the system. Programmers often used space saving short-cuts for field names, also causing confusion. Often it is only the developers, who may have long since moved on, who know what a name means or refers to, or how it relates to other fields.

If you want a particular piece of data, it can be hard to know where to go in the system to get it. So of those 100,000 or so fields within an operational system, it can be difficult for a bank working on its own to pick out everything it needs to know.

Some companies undertaking data warehouse projects use homegrown code to support ETL processes. However, even with successful implementations, it is likely the source data file formats, and the validation rules applying to the data, will evolve. This means regular modification and maintenance of the ETL code. Companies can encounter further problems when adding systems or as the volume of data grows.

This does not mean it is impossible for a financial institution to go it alone with ETL. At West Bromwich, Eggleston uses a system that is 'mostly our own work'. He says, as a building society, West Bromwich doesn't have a complex set of data. Moreover, buying in a third party solution would have been unjustifiably expensive. Even so, a consultancy was called in for the initial stages.

As a generic tool, (there are many versions available) ETL does pretty much what it says. The Extract part of ETL is an automated process. Generally, data is retrieved from the various operational systems using read-only native SQL. SQL directly queries the client's data within the core system. Interfering with the core system directly may compromise the integrity of the system, and is also likely to contravene the software licence agreement

## BI: Potential for external reporting?

Although in theory a BI system that generates audit trails could be used for external reporting duty, for the moment at least this is unlikely as central bank system-approval is required. But, says Richard Green at Cooperative Bank Treasury, 'System-approval from the central bank tends only to relate to the data format. The audit trail is up to the organisation.' To comply, most banks have had to put a link in somewhere between their general ledger and the final report form. This is usually done at a higher point than transaction level. The ledger extract is put into a central bank report, but it will often require some kind of customer analysis too. The latter data is unlikely to be stored on the ledger, so data will have to come from an underlying system. This has to be controlled to the ledger before it is fed into the central bank reporting suite: You can't rely totally on the underlying transaction system. 'If you have loaded your ledger and transaction system into a relational database, you can do that control in the database, and then just have the reporting software format and post it as the final stage,' says Green. So, rather than replacing a proprietary reporting package, it may be beneficial to run a BI system alongside it, to assist and enhance compliance. There are those of course who would rather leave the external reporting function to the specialist vendors. Achim Vogt, COO of Deutsche Postbank, London, is one. He offers a simple explanation: 'Reporting rules change too rapidly. We don't want to know how to do it in such detail.'

of the core system vendor. For this reason it is highly likely that the SQL system a vendor offers will need to be fully audited by the client, and signed off as acceptable.

In Transformation mode, data extracted from the source system is automatically 'cleaned' (de-duplicated and reconciled against all the other primary sources of data), and mapped to the data model. The simple automated process of loading it into the new data warehouse is then executed. What a bank ends up with is a warehouse full of essential metadata primed for BI.

Having populated the warehouse, the ETL process to refresh data is generally operated to a schedule, normally after each end-of-day. It can also be delivered programatically at any point, or worked ad hoc, in real-time, by the user. All financial data must be reconciled back to the general ledger, otherwise it has no credibility.

Providers of packaged ETL systems include Microsoft, which offers data transformation services bundled with its SQL Server database. Oracle has embedded ETL capabilities in its database, and IBM offers a DB2 Information Integrator component for its warehouse offerings. Specialist BI vendors, and those offering third party solutions for data integration, may be able to handle a broader spectrum

of source system applications and data structures than these proprietary offerings.

Once data is in the warehouse, you need to make it work for its keep. To go further into it to make comparisons and deliver alternative views can bring into play a multitude of specialised BI applications at the delivery stage: scorecards, data mining, query and reporting, business process flow analysis, business performance management, text mining, CRM and so on.

Some vendors may offer a complete package containing their own variations on the range of applications, whilst others may encourage the flexibility of a best of breed component-based approach. Peddar espouses the latter. 'You are future proofing much more. If someone comes along with a better tool, you can swap it - provided they can work together within an industry standard.' And levels of interoperability are vastly improved, he believes.

Business Objects' Neale looks to the single vendor platform as the best way. He argues that best-of-breed equals expensive and time-consuming. Costly operational downtime arises from unplugging the product and interfacing the new. Integration could be a problem, and extra training is required on each new product, meaning staff inefficiencies until they are up to speed, he says. He also suggests that

BANKING TECHNOLOGY. *Redefined.*



SCALABLE

SEAMLESSLY INTEGRATED

FRONT TO BACK CRUISE-STEP

RETAIL & WHOLESALE BANKING COVERAGE

NEW GENERATION TECHNOLOGY PLATFORM

ADAPTABLE INFORMATION ARCHITECTURE

**SYMBOLS**  
UNIVERSAL BANKING SOLUTION



www.systemaccess.com  
Singapore Bratislava Bangkok  
Dubai Manila Prague

## EXAMPLE BI SOLUTION PROVIDERS

| COMPANY                  | SYSTEM                       | WEBSITE                               | EXAMPLE CLIENTS  |
|--------------------------|------------------------------|---------------------------------------|--|
| Applix Inc               | ApplixTMI                    | www.applix.com                        | ING BHF Bank   |
| Business Objects         | BusinessObjectsXI            | www.businessobjects.com               | Claims 29,000 in 80 countries  |
| Cognos                   | Cognos Business Intelligence | www.cognos.com                        | Regions Financial, Bank of Ireland, Key Corp, Bear Stearns, Citigroup, UBS                   |
| Granville Associates     | Granville IM                 | www.granvilleassociates.co.uk         | Deutsche Postbank London   |
| Hummingbird              | HummingbirdBI                | www.hummingbird.com                   | No financial clients listed  |
| Hyperion (formerly Brio) | Hyperion Intelligence        | www.brio.com                          | Amsouth Bank, Banca 24-7, Citibank, Banca National de Costa Rica, Bank One, HBOS, Woori Bank |
| Information Builders     | WebFocus                     | www.informationbuilders.com           | PNC Bank, RBC Financial Group, Bank Sarasin Benelux, BBVA                                    |
| Information Mgt Group    | IMGROUP BI                   | www.imgroup.com                       | No financial clients listed  |
| In Group                 | inGroup BCI                  | www.ingroup.co.uk                     | Lloyds TSB, Caledonia Investments, GE Insurance  |
| Micro Strategy           | MicroStrategy BI Platform    | www.microstrategy.com                 | Bank of Montreal, Bank Rakyat, Export-Import Bank of Korea, FleetBoston Financial            |
| Netezza                  | Netezza Performance Server   | www.netezza.com                       | Orange, AT&T Wireless, Epsilon   |
| Oracle                   | Oracle Application Server    | www.oracle.com/technology/products/bi | Banco de Credito, Abbey National, Commerzbank, Shell International                           |
| Siebel                   | Siebel Business Analytics    | www.siebel.com                        | Bank of America, Société Générale, American Express  |

each new product needs to be managed and administered separately from the core BI system and any application modifications can also be costly and difficult. Cost of upgrades and new releases for each component also need consideration.

Despite the plethora of front-end applications, Excel is perhaps the most commonly used. As Excel is a simple two dimensional tool, interfacing to the BI platform brings a quick and easy way of switching between dimensions, or drilling down into data. Bi-directional functionality also allows reconciliation between spreadsheets, formulae validation, and all data changes to be verified and audited, sometimes right down to cell level.

### OLAP tools and variations

One way of achieving far greater Excel flexibility is through an On-Line Analytical Processing (OLAP) application, for which there are a number of major suppliers such as Microsoft, IBM, Cognos, Business Objects, and Applix.

OLAP has been designed to provide answers to complex database queries in the shortest possible time. The chief component of an OLAP application is the OLAP server which sits between a client and a database. There are OLAP servers available for nearly all the major database systems. Delivered from a single server, the report format is usually based on common XML, so reports can be opened in most client suites.

OLAP technology also requires a data storage component and a query processing engine. A number of data storage variations for OLAP exist. In a relational database version, the technology is known as ROLAP. Here, storage of multi-dimensional OLAP data is usually achieved by having a fact table and a dimension table stored under a star schema (in a relational database, the schema defines the tables, the fields in each table, and the relationships between fields and tables). Each of these tables, called a cuboid, represents a particular view. A ROLAP engine can use either a multi-pass

SQL or a multi-dimensional processing engine (MPE) for processing. The MPE can be located on the server or at the client. As the cuboids are conventional database tables, they can be processed and queried using traditional RDBMS techniques, such as indexes and joins.

A ROLAP engine uses relational data tables directly to produce web reports and data analysis on request. The request is converted into one or more SQL statements that are submitted to the relational sources. These respond with result tables. The results are then combined and formatted for the user.

ROLAP has the advantage of always supplying completely up to date data if it is refreshed regularly via the ETL programme. However, relational databases offer relatively slow performance when looking at a large amount of transactional data. Also, ROLAP solutions are only available to remote users as client-server applications because it is impractical to shift a large database to the end-user. ROLAP tools are



Richard Peddar,  
Granville Associates

best employed by users who often have an unpredictable set of queries.

Another storage option is a multi-dimensional database. This is often referred to as an OLAP cube or Hypercube. Multi-dimensional 'cubes' of data can automatically be constructed from summaries of relational tables, based on the users' required cube definitions. The engine uses the cubes to compute results. Extensive pre-calculation and aggregation is possible in this storage scheme and, because the data is summarised, it allows much faster query processing. Multiple cubes may be created and linked, so if an element that is common between cubes is changed in one dimension, it changes in all, the system automatically re-calculating. This implementation is known as MOLAP. MOLAP tools are better suited to users with habitual query sets (they ask the same questions every day/week/month on an updated cube).

HOLAP (Hybrid) technology is an attempt to combine the advantages of MOLAP and ROLAP. For summary-type information, HOLAP uses cubes for rapid querying. But when detailed information is required, HOLAP can jump out of the cube into the underlying relational data. Oracle's 10g offering is one of the first to offer this combination.

Cubes (a notional title: effectively you can have as many 'sides', or dimensions, as required) act like windows on data within spreadsheets and allow very fast access to

your data summaries. You can also move from window to window, enabling a different view from each, and users can drill down into the data to view at a more granular level - a process known as 'slicing and dicing'.

Data aggregation means cubes are essentially compressed versions of your data and are quick and easy to distribute. The downside of MOLAP is that it doesn't scale to large data sets as well as ROLAP. Another issue, according to Neale, is that a cube's data calculations are fixed. 'It is frozen in terms of the dimensions you want to look at. New cubes need to be built for each extra dimension required.'

MOLAP data only contains aggregations of raw data up to the level where records are described in the cube's virtual space. If data is required for all transactions in one day, the cube can only give an aggregated figure if its date dimension only distinguishes to the day. Further aggregations can be stored in a cube, allowing the user to break the data down, but the larger the data set, the higher the number of intermediate sums a cube must contain. The number of combinations of dimensions and members of the data set grow exponentially with the numbers of those elements in the cube - the result can be an undesirable 'data explosion'. Neale believes cubes are best suited to tactical analysis, to produce quick data for short-term projects where response times are paramount, leaving large data sets to ROLAP applications.

But, argues Peddar, cubes are quick and easy to build, re-build, and throw away. 'You can build a simple cube to support simple data, then build it up, adding more information and dimensions.' Cubes can be added or taken away at any point; indeed, some users may have access to build their own cubes under a developer licence.

#### Processes... and people

The underlying technology may have come a long way in the last ten years or so, but getting impressive results is about more than simply plugging a BI platform into an operational system. Last year Alan Pendse, in his third influential Survey.com report on OLAP technology, tells how, 'on the whole, product failings are less of a problem than fractious peoples'.

A condition of implementing a BI platform (or any other major IT solution) should be the simultaneous instigation of an organisational change management programme. This is part of the project management function. It is crucial for an organisation to recognise that processes *and* people make for success. Software solutions on their own, without a carefully planned and managed architecture, are simply not solutions. 'You need architects to design and understand where you are going, and keep changing it in line with how the business evolves,' says Peddar. 'Unless that architecture is in place, platforms like BI can't be optimised.'

If a project lasts much longer than six months, problems arise. 'People move around too much, they get bored. Banks want short-term tangible benefits. If it's bigger than that, break it down into multiple projects,' says Peddar. He also stresses the importance of using high-potential internal staff with clear views and leadership skills. Before the first project is finished, he suggests, set up the next two - and keep growing. This must be carried out within the structure of the overall architecture, but flexible planning is key in the long-term. Short-term goals within the framework will succeed. A hard ten-year plan will fail, he says. Neale agrees: 'You're looking for an iterative type project. Look for a quick win that can get the momentum of the project going.' In short, tools alone will not do the job.

The technology seems to have reached a point now where it can deliver. Certainly Eggleston is happy with the West Bromwich system: 'It is always edifying to look back at the original project case and see now that over 90 per cent of what we set out to achieve has been delivered.' In the quest for an easy to use point and click product, he says: 'The system is acting exactly as we expected.'

There are a number of reasons why BI can be a useful addition for the many, rather than an expensive option for the few. But there are clearly a number of caveats for banks thinking of taking this route. If you were looking for a fundamental rule, a good place to start would be Eggleston's suggestion to, 'Just make it as simple as possible'. 